

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 07:04:57 ; Search time 9694.27 Seconds
(without alignments)
12038.564 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGAAACGACATGACCATGA.....GGTACCCAAATCGCCCTATA 3066

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hcc:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gsl1:*

9: gb_gsl2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2342.4	76.4	2610	3	BC024044 Homo sapi
2	809.4	26.4	1086	5	BM912146 AGENCOURT
3	788.8	25.7	1038	2	BE249939 600942990
4	753.2	24.6	837	4	BG290366 602388204
5	737.8	24.1	1089	5	EX419743 BX419743
6	696.4	22.7	897	7	CN646528 ILLUMIGEN
7	595.8	19.4	5738	3	CR749643 Homo sapi
8	595.6	19.4	2339	3	AK008957 Mus muscu
9	594.8	19.4	630	1	AL120048 DKEZp761K
10	592.6	19.3	820	7	CO724689 ILLUMIGEN
11	586.4	19.1	1076	4	BG702473 602684537
12	583.4	19.0	598	5	BP199106 BP199106
13	581.4	19.0	583	5	BP342407 BP342407
14	581	18.9	589	4	BM702087 UI-B-CQ1-
15	573.2	18.7	583	5	BX486692 DKEZp686P
16	566.8	18.5	583	5	BP195050 BP195050
17	564.2	18.4	582	5	BP193074 BP193074
18	561.6	18.3	723	7	CN647268 ILLUMIGEN
19	561.2	18.3	578	5	BP341239 BP341239
20	560.6	18.3	581	5	BP216015 BP216015
21	557.4	18.2	588	5	BP193198 BP193198
22	537.2	17.5	581	5	BP326142 BP326142
23	536.8	17.5	677	4	BP549901 603194747
24	532	17.4	564	5	BP221049 BP221049

25	530.6	17.3	579	5	BP342645	BP342645
26	520.8	17.0	710	5	BP383205	BP383205
27	519.2	16.9	563	5	BP221191	BP221191
28	515.6	16.8	765	7	CO725626	CO725626
29	512	16.7	582	5	BP213083	BP213083
30	510	16.6	668	4	BG468930	BG468930
31	506	16.5	676	7	CK002754	CK002754
32	497.2	16.2	583	5	BP195817	BP195817
33	496.8	16.2	582	5	BP308450	BP308450
34	496	16.2	508	2	BE293603	BE293603
35	495	16.1	496	1	AI573188	AI573188
36	494.2	16.1	598	5	BP217439	BP217439
37	490.6	16.0	965	4	BG261180	BG261180
38	488	15.9	588	5	BP310687	BP310687
39	486	15.9	490	5	EX282664	EX282664
40	482.4	15.7	787	2	BF675198	BF675198
41	475.8	15.5	773	5	BX463597	BX463597
42	474.8	15.5	497	6	CD617341	CD617341
43	474.8	15.5	599	2	BE866183	BE866183
44	469.2	15.3	581	5	BP194681	BP194681
45	469	15.3	1104	7	CN643641	CN643641

ALIGNMENTS

RESULT 1
LOCUS BC024044 2610 bp mRNA linear HTC 23-SEP-2002
DEFINITION Homo sapiens, Similar to band 83, clone IMAGE:4645428, mRNA.
ACCESSION BC024044
VERSION BC024044.1 GI:23270792
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2610)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (11-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgaps-re@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natagja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL plate: 29 Row: m Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4502372
This clone has the following problem: frame shifted.

FEATURES
source
1. .2610
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:4645428"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC 15"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 76.4%; Score 2342.4; DB 3; Length 2610;
Best Local Similarity 91.7%; Pred. No. 0;
Matches 2598; Conservative 0; Mismatches 1; Indels 235; Gaps 2;

Qy	135	GCTCCACGACGACCGACGCTGGGCGACGACGCTGAGAGACCCAGGACCCCTGTGCAGGA	194
Db	1	GCTCCACGACGACCGACGCTGGGCGACGACGCTGAGAGACCCAGGACCCCTGTGCAGGA	60
Qy	195	GCAGTCCGGGTGACACGAGGGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAAT	254
Db	61	GCAGTCCGGGTGACACGAGGGGACTGAAGATACCTCCACAGGGGCTCAGCAGGAGCAAT	120
Qy	255	GGGTAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGA	314
Db	121	GGGTAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGA	180
Qy	315	GACTTACAGGACAAACGGCTGTCTGAAACCGGGTTCAGTGGTGGTTCGACCCACAC	374
Db	181	GACTTACAGGACAAACGGCTGTCTGAAACCGGGTTCAGTGGTGGTTCGACCCACAC	240
Qy	375	AGTTTCAGACATAGAGGAGTGCCTTGGGATAGTGTCAAGCGGATATGTGGCCAC	434
Db	241	AGTTTCAGACATAGAGGAGTGCCTTGGGATAGTGTCAAGCGGATATGTGGCCAC	300
Qy	435	TTCTTCCCGGAGCAACCGGAGTAAGTGTCTGCGGATGCCAACGGAAAGAAATCTTGG	494
Db	301	TTCTTCCCGGAGCAACCGGAGTAAGTGTCTGCGGATGCCAACGGAAAGAAATCTTGG	360
Qy	495	GAAGAGGCGCAAAACCGGAGGACGAGTGTCTGAAACCGGGTTCAGTGGTGGTCTCTCG	554
Db	361	GAAGAGGCGCAAAACCGGAGGACGAGTGTCTGAAACCGGGTTCAGTGGTGGTCTCTCG	420
Qy	555	GCCTGTACGAGCGTACCGGAGCCAGCCAGCGCGAGATTCATCCCTTGGATCAGTGAAGCT	614
Db	421	GCCTGTACGAGCGTACCGGAGCCAGCCAGCGCGAGATTCATCCCTTGGATCAGTGAAGCT	480
Qy	615	TGATGTACGCTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAAGCTGCACTTCC	674
Db	481	TGATGTACGCTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAAGCTGCACTTCC	540
Qy	675	GGTGCAGCTGGAACCGGGGAGGACACAGATAAAACCCCGAGGCAAGCCCGGCCCAAGA	734
Db	541	GGTGCAGCTGGAACCGGGGAGGACACAGATAAAACCCCGAGGCAAGCCCGGCCCAAGA	600
Qy	735	CAAGTCTCTCTGCGCCAGGGATCCACCGCTTCTCCAGCTGAGACAGAGGGGAGCAGG	794
Db	601	CAAGTCTCTCTGCGCCAGGGATCCACCGCTTCTCCAGCTGAGACAGAGGGGAGCAGG	660
Qy	795	AGGGAAGCTCCCTCCAAAGCCCAAGGACTCCAGC- TTTTTCACAAATTTCTTCAAGCTGG	853
Db	661	AGGGAAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTTCACAAATTTCTTCAAGCTGG	720
Qy	854	ACAAGGGAACGAAAGGTGACAGGTGACAGCCAAACAGGAAACCAAGAGGGGAGCAATC	913
Db	721	ACAAGGGAACGAAAGGTGACAGGTGACAGCCAAACAGGAAACCAAGAGGGGAGCAATC	780
Qy	914	AAGCAAGGTGGATGAGTCTCTGCTTATCAGGGCAGTCCGATGATGTCCTCGAGGA	973
Db	781	AAGCAAGGTGGATGAGTCTCTGCTTATCAGGGCAGTCCGATGATGTCCTCGAGGA	840
Qy	974	AGGACATAGTTGACCGGCAAGGAAAGAGGACAGAACTTGGAACTCGGATGCTCTG	1033
Db	841	AGGACATAGTTGACCGGCAAGGAAAGAGGACAGAACTTGGAACTCGGATGCTCTG	900
Qy	1034	TCCCTGGGACCCAGAAAGCTGGAGACTGCAAGGACGATTCACAGGACGATGATATAG	1093

Db	901	TCCCTGGGACCCAGAAAGGACTGGAGACTGCACAAAGSACGATTCCTCCAGGACGACTATAG	960
Qy	1094	CAGAGATAATAATTTCCATCATGAGTTTCTTTAAACTCTCTGGTTTCACTTAAACAAAGCTG	1153
Db	961	CAGAGATAATAATTTCCATCATGAGTTTCTTTAAACTCTCTGGTTTCACTTAAACAAAGCTG	1020
Qy	1154	AAACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAAGTCAACCCACCTTTCAGCTGACC	1213
Db	1021	AAACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAAGTCAACCCACCTTTCAGCTGACC	1080
Qy	1214	TTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAGAAATTTCCA	1273
Db	1081	TTAAGTCAGACAAAGCCAACTTTTACATCCAGGAGACCCAAAGGGGCTGGCAGAAATTTCCA	1140
Qy	1274	AAGGATCAACCCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCTCGGAAGGAAG	1333
Db	1141	AAGGATCAACCCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACTCTCGGAAGGAAG	1200
Qy	1334	GCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGT	1393
Db	1201	GCACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGT	1260
Qy	1394	CAGTTAAAGAGGACTCAGTCCCAACAGTGGGAGGAGAAATGCTGGTGTGAGTCAACAG	1453
Db	1261	CAGTTAAAGAGGACTCAGTCCCAACAGTGGGAGGAGAAATGCTGGTGTGAGTCAACAG	1301
Qy	1454	TAGAGATTATAAAGTCCAAGGAAGTAGAATCAGCCTTTACAAACAGTGGACCTCAACGAAG	1513
Db	1302	-----	1301
Qy	1514	GAGATGTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAAAACCAAGAA	1573
Db	1302	-----	1301
Qy	1574	CCTCTGTAGTGGCGTTTCTCAGACAAATGTCTAGTGAAGGGGATGGAGGGATCACCCT	1633
Db	1302	-----	1301
Qy	1634	CAGAAGAAATAAATGGGAAGAGCTCCAGTGCCTCAACATCAGACTCCACAGAAAGACTA	1693
Db	1302	-----	1301
Qy	1694	TCACACCGCCAGAGCTTGAACCAACAGGAGCAACACAGAGGGTAAAGAGGGCTCTCTGA	1753
Db	1327	TCACACCGCCAGAGCTTGAACCAACAGGAGCACACAGAGGGTAAAGAGGGCTCTCTGA	1386
Qy	1754	AGGACAAAGAGTCTCAGACGCGGAGATGAACAGCAGAGAGAGCAACAGCAGAGAGCCAAAG	1813
Db	1387	AGGACAAAGAGTCTCAGACGCGGAGATGAACAGCAGAGAGAGCAACAGCAGAGAGCCAAAG	1446
Qy	1814	AACAGGCGCAGTGCACAGAGCAGGCGCCACGGTGGACACGAACTCACTGCAGAAATGGGACA	1873
Db	1447	AACAGGCGCAGTGCACAGAGCAGGCGCCACGGTGGACACGAACTCACTGCAGAAATGGGACA	1506
Qy	1874	AGCTCCAAAAGAGACCTGAGAAAGCGGAGCAGTCCCTTTGGGGGCTTCTTTAAAGGCTGG	1933
Db	1507	AGCTCCAAAAGAGACCTGAGAAAGCGGAGCAGTCCCTTTGGGGGCTTCTTTAAAGGCTGG	1566
Qy	1934	GACCAAGCCGATGTTGGATGCTCAAGTGCNAACAGACCCAGTATCCATCCGACCAAGATGG	1993
Db	1567	GACCAAGCCGATGTTGGATGCTCAAGTGCNAACAGACCCAGTATCCATCCGACCAAGATGG	1626
Qy	1994	GCAAAACCAAGTAAACAAATCAGCACGCTTCCACAGGTTCTCTGCTCCACCAAGATGTG	2053
Db	1627	GCAAAACCAAGTAAACAAATCAGCACGCTTCCACAGGTTCTCTGCTCCACCAAGATGTG	1686
Qy	2054	TTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATTTCTTGTAGTGGCCAG	2113
Db	1687	TTCTCTTACTTCCATCTCTCTCCCAACACGCTCCATGTATATATTTCTTGTAGTGGCCAG	1746
Qy	2114	CAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGGCTGTATTTATTA	2173
Db	1747	CAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTTAGGCTGTATTTATTA	1806

QY 2174 CGTCTCTGGTCCAGTCTTTCTGGCAATAACAGTAAGATGGTTTACGAGTCACTAG 2233
DB 1807 CGTCTCTGGTCCAGTCTTTCTGGCAATAACAGTAAGATGGTTTACGAGTCACTAG 1866
QY 2234 TTGGGTTCAGAGAGTCCGATGATCACCAGCAGAAAGGAGGCAATAGAGGAATGTGTTTC 2293
DB 1867 TTGGGTTCAGAGAGTCCGATGATCACCAGCAGAAAGGAGGCAATAGAGGAATGTGTTTC 1926
QY 2294 GGGTTAAGTGAATAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCCCTGTGAATCTCA 2353
DB 1927 GGGTTAAGTGAATAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCCCTGTGAATCTCA 1986
QY 2354 GCATTTTCGGAGCGCGAGTGGATCCTGAGGTTCAGAGTTCAGACTAGCCTGG 2413
DB 1987 GCATTTTCGGAGCGCGAGTGGATCCTGAGGTTCAGAGTTCAGACTAGCCTGG 2046
QY 2414 CCAACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGCATGGTGGCAC 2473
DB 2047 CCAACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGCATGGTGGCAC 2106
QY 2474 ACACCTGTAGTCCAGTCTCTGAGGAGCCCAACGACGAGAACCGCTTTGTACCCAGGAGG 2533
DB 2107 ACACCTGTAGTCCAGTCTCTGAGGAGCCCAACGACGAGAACCGCTTTGTACCCAGGAGG 2166
QY 2534 TGGAGTTTCAGTGAAGCCGAGTTGCACATTCGACTCCACCTGGGCGACAGAGCAAGA 2593
DB 2167 TGGAGTTTCAGTGAAGCCGAGTTGCACATTCGACTCCACCTGGGCGACAGAGCAAGA 2226
QY 2594 TTCTATCAAAAAAAGGCGAGTGCAGTGAAGTTATAGAGAGAAATGCTGTAGAGG 2653
DB 2227 TTCTATCAAAAAAAGGCGAGTGCAGTGAAGTTATAGAGAGAAATGCTGTAGAGG 2286
QY 2654 AATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGAGAGCAAAAT 2713
DB 2287 AATTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTGAAGAGGAGAGCAAAAT 2346
QY 2714 CCAATTTGTTTAAATTCATCTCAAGGAGGAGAAACCGGGGTGTGTGGGTGTGTTGCCA 2773
DB 2347 CCAATTTGTTTAAATTCATCTCAAGGAGGAGAAACCGGGGTGTGTGGGTGTGTTGCCA 2406
QY 2774 ATTTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAA 2833
DB 2407 ATTTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAA 2466
QY 2834 ATAGGTCCTTTGTAAAGTTATTCAGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATG 2893
DB 2467 ATAGGTCCTTTGTAAAGTTATTCAGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATG 2526
QY 2894 ATTTGGGAAAAACAATTTGTTTGGGCTCAGTGACACCGGCAAGATTACAACTTAAAAA 2953
DB 2527 ATTTGGGAAAAACAATTTGTTTGGGCTCAGTGACACCGGCAAGATTACAACTTAAAAA 2586
QY 2954 AAAAAAATAAAAAA 2967
DB 2587 AAAAAAATAAAAAA 2600

RESULT 2
BM912146
LOCUS BM912146
DEFINITION AGENCOURT_6613041 NTH_MGC_41 Homo sapiens cDNA clone IMAGE:5473643
5', mRNA sequence.
ACCESSION BM912146
VERSION BM912146.1 GI:19362525
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1086)
NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: DCTD/BTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCMI985 row: j column: 12
High quality sequence start: 13
High quality sequence stop: 510.

FEATURES
source

Location/Qualifiers
1..1086
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5473643"
/tissue_type="amelanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_41"
/note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN

Query Match 26.4%; Score 809.4; DB 5; Length 1086;
Best Local Similarity 95.4%; Pred. No. 3.3e-121;
Matches 867; Conservative 0; Mismatches 36; Indels 6; Gaps 3;
QY 279 AAGAGTTGAACACCAAGAGATGAACCAAGAGAGAGACTTACCAAGGACAACCGCTCTGC 338
DB 1 AAGAGTTGAACACCAAGAGATGAACCAAGAGAGAGACTTACCAAGGACAACCGCTCTGC 60
QY 339 TCTCAACGGGGTTCAGTGGTGTGTCGACCCACACAGTTCAGCACTTAGAGGAAGTGA 398
DB 61 TCTCAACGGGGTTCAGTGGTGTGTCGACCCACACAGTTCAGCACTTAGAGGAAGTGA 120
QY 399 CTTCGGGAATAAGTGTCAAGACGGAATAATGTGGCCAATTTCTTCCCCCGAGACAACGGAGAT 458
DB 121 CTTCGGGAATAAGTGTCAAGACGGAATAATGTGGCCAATTTCTTCCCCCGAGACAACGGAGAT 180
QY 459 AAGTGTCTGTTGGGATGCGCAACGGAAGAAATCTTGGGAAAGAGGCCAAACCGAGGACCC 518
DB 181 AAGTGTCTGTTGGGATGCGCAACGGAAGAAATCTTGGGAAAGAGGCCAAACCGAGGACCC 240
QY 519 AGCTGTCTAAATCTCGTTTTTTTCTGATGCTCTCTCGGCTGTACCAAGCTTCAATAAAGCTCC 578
DB 241 AGCTGTCTAAATCTCGTTTTTTTCTGATGCTCTCTCGGCTGTACCAAGCTTCAATAAAGCTCC 300
QY 579 CCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTTGTGATGCTCAGTCCAATAAAGCTCC 638
DB 301 CCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTTGTGATGCTCAGTCCAATAAAGCTCC 360
QY 639 AGCGAACAAGACCAAGTGAAGTGAACACTTCCGGTGGCAGCTGAGCGGGGACAGA 698
DB 361 AGCGAACAAGACCAAGTGAAGTGAACACTTCCGGTGGCAGCTGAGCGGGGACAGA 420
QY 699 CACAGATAAACCACCGGCAACCGCCCGCCAGAGCAAGGTCTCTCTCGGCGCAGGA 758
DB 421 CACAGATAAACCACCGGCAACCGCCCGCCAGAGCAAGGTCTCTCTCGGCGCAGGA 480
QY 759 TCCCACTGCTTCTCCCACTGAGACAGGGGGAGCAGGAGAGAAGCTCTCCCTCAAGCCCCAA 818
DB 481 TCCCACTGCTTCTCCCACTGAGACAGGGGGAGCAGGAGAGAAGCTCTCCCTCAAGCCCCAA 540
QY 819 GGACTCCAGCTTTTTTTTGACAAAATTTCTTCAAGCTGGACAAGGACAGGAAAAGGTGCCAGG 878

```
Db 541 GGACTCCAGCTTTTGTGACAAATTCCTCAAGCTGGACAAAGGCACAGGAAAAGGTGCCAGG 600
Qy TGACAGCCAAACAGGAAGCCAAAGAGCGGACAGCATCAAGACAAAGTGGATGAGTTCCTGG 938
Db 601 TGACAGCCAAACAGGAAGCCAAAGAGCGGACAGCATCAAGACAAAGTGGATGAGTTCCTGG 660
Qy 939 CTTATCAGGGGAGTCCCGATGATGTCCTCGCAGGGAAGGACATAGTTGACGGCAAGGAAA 998
Db 661 CTTATCAGGGGAGTCCCGATGATGTCCTCGCAGGGAAGGACATAGTTGACGGCCACGAGA 720
Qy 999 AGAAGGACAAAGACTTGGAATCGCGGATGTCCTCGCTGGGACCCAGAACGACTGGA 1058
Db 721 CGAAGGACCAAGACTTGGAATCGCGGATGTCCTCGCTGGGACCCAGAACGACTGGA 780
Qy 1059 GACTGCAAGGACGATTCCTCAGGAGGAGCTATAGCAGAG--AATAAATATTCATCATGA 1117
Db 781 GACTGCCAAGGACGATTCCTCAGGAGGAGTATAGCCAGAAATATATATTTCCATGCTGA 840
Qy 1118 GTTTCTTTT--AAAACCTCTGGTTTCACTTAACAAAGCT---GAAACAAAAGGACCCAGA 1172
Db 841 GTTTCTTTTAAACTCTGGTTTCACTTACAAAGCTTGAGAACCAAGGAAACCCGAA 900
Qy 1173 AGACACGGG 1181
Db 901 AAAAACCGG 909
```

```
RESULT 3
BE249939
LOCUS 600942990F1 NIH_MGC_15 Homo sapiens cDNA clone IMAGE:2959439 5',
DEFINITION BE249939 1038 bp mRNA linear EST 13-JUL-2000
RNA sequence.
ACCESSION BE249939
VERSION BE249939.1 GI:9120042
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1038)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: sgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW51 row: c column: 24
High quality sequence stop: 784.
Location/Qualifiers
1..1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:2959439"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
```

```
FEATURES
source
1..1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="caxon:9606"
/clone="IMAGE:2959439"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_15"
/notes="Organ: colon; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dr priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)"
```

ORIGIN

```
Query Match 25.7%; Score 788.8; DB 2; Length 1038;
Best Local Similarity 94.5%; Pred. No. 7,2e-118;
Matches 862; Conservative 0; Mismatches 42; Indels 8; Gaps 4;
Qy 656 GTGAGAGCTGGACACTTCCGGTGCACCTGGACCGGGCAGGACACAGATAAACCACCG 715
Db 1 GTGAGAGCTGGACACTTCCGGTGCACCTGGACCGGGCAGGACACAGATAAACCACCG 60
Qy 716 GGACAGCCCCGGCCAAAGACAAGGTCTCTCTGCGCCAGGGATCCCAACGCTTCTCCAC 775
Db 61 GGACAGCCCCGGCCAAAGACAAGGTCTCTCTGCGCCAGGGATCCCAACGCTTCTCCAC 120
Qy 776 CTGAGACAGGGGGAGCAGGAGGAAGCTCCCTCCAAAGCCAAAGACTCCAGCTTTTGT 835
Db 121 CTGAGACAGGGGGAGCAGGAGGAAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGT 180
Qy 836 ACAATTTCTTCAAGCTTGGACAAAGGACAGGAAGGTGCCAGTGCAGAGCCAAACAGAG 895
Db 181 ACAATTTCTTCAAGCTTGGACAAAGGACAGGAAGGTGCCAGTGCAGAGCCAAACAGAG 240
Qy 896 CCAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCG 955
Db 241 CCAAGAGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCG 300
Qy 956 ATGATGTCTCTGACAGGAGGACATAGTTGACGCAAGGAAAAAGAGGACAAAGACTTG 1015
Db 301 ATGATGTCTCTGACAGGAGGACATAGTTGACGCAAGGAAAAAGAGGACAAAGACTTG 360
Qy 1016 GAACCTGGGATGCTCTGCTCCCTGGGACCCAGAGAGGACTGGAGACTGCAGAGGACGAT 1075
Db 361 GAACCTGGGATGCTCTGCTCCCTGGGACCCAGAGAGGACTGGAGACTGCAGAGGACGAT 420
Qy 1076 CCCAGGACAGAGCTATAGCAGAGAATAAATTCATCATGAGTTCCTTTAAAACTCTGG 1135
Db 421 CCCAGGACAGAGCTATAGCAGAGAATAAATTCATCATGAGTTCCTTTAAAACTCTGG 480
Qy 1136 TTTCACTTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACGGGTGCTGAAAGTCAAC 1195
Db 481 TTTCACTTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACGGGTGCTGAAAGTCAAC 540
Qy 1196 CCACCACTTCAAGCTTAAAGTCAGACAAAGCCAACTTTACATCCAGAGAGACCCAG 1255
Db 541 CCACCACTTCAAGCTTAAAGTCAGACAAAGCCAACTTTACATCCAGAGAGACCCAG 600
Qy 1256 GGGCTGCAAGAAATTCACAAAGGATGCAACCATCCGGGCACACACAGTCCCTGCAACCC 1315
Db 601 GGGCTGCAAGAAATTCACAAAGGATGCAACCATCCGGGCACACACAGTCCCTGCAACCC 660
Qy 1316 CTGAACCTGCAAGAGGACCAAGGAGAAATCAGACCCACTCTCTGCTCTGGGCA 1375
Db 661 CTGAACCTGCAAGAGGACCAAGGAGAAATCAGACCCACTCTCTGCTCTGGGCA 720
Qy 1376 AACTGTTTGGAAAAAGTCAGTTAAGAGGACTCAGTCCCCACACAGGTG-CGGAGGAGAT 1434
Db 721 AACTGTTTGGAAAAAGTCAGTTAAGAGGACTCAGTCCCCACACAGGTGCGCGGGAGAT 780
Qy 1435 GTGGTG-TGTGAGTCACCAAGTAGAGATTATAAGTCCAAAGGAAGTAGAATCAGCTTACA 1493
Db 781 GTGGTGTTGTGAGTCCCAAGTAGAGAT--TTAAGTCCCGCCAGTTGATCTCGCTACAA 838
Qy 1494 AACAGTGGACCTCAACGAAGGAGATGCTGCACCTGAAACCCACAGAGGAGAACTCAAAAG 1553
Db 839 CCAGGGGGCACTCACCGGGA---ATGCGGCTCTGAACCCCAAGAGCGAATCTCAAGAG 894
Qy 1554 AGAAGAAAGCAA 1565
Db 895 AAAAGGCAACCA 906
RESULT 4
BE290366
LOCUS 602388204F1 NIH_MGC_93 Homo sapiens cDNA clone IMAGE:4517041 5',
DEFINITION
```



```

mRNA sequence.
ACCESSION BG290366
VERSION BG290366.1 GI:13047155
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL NIH-MGC http://mgc.nci.nih.gov/.
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: L1AM10409 row: h column: 02
High quality sequence stop: 718.
FEATURES             source
   1..837
    Location/Qualifiers
        1..837
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:4517041"
            /lab_host="DH10B (phage-resistant)"
            /clone_lib="NIH MGC_93"
            /note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;
            Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
            Average insert size 1.7 kb. Library enriched for
            full-length clones and constructed by Life Technologies.
            Note: this is a NIH_MGC Library."
ORIGIN
Query Match      24.6%; Score 753.2; DB 4; Length 837;
Best Local Similarity 97.5%; Pred. No. 4.4e-112;
Matches 786; Conservative 0; Mismatches 18; Indels 2; Gaps 2;
QY 1142 CTAACAAGCTGAACAAAGGACCCAGAGACACCGGTGCTGAAAGTCAACCCACCA 1201
DB 1 CTAACAAGCTGAACAAAGGACCCAGAGACACCGGTGCTGAAAGTCAACCCACCA 60
QY 1202 CTTGAGCTGACCTTAAGTCAGACAAAGCCAACTTTACATCCAGGACCCCAAGGGGCTG 1261
DB 61 CTTGAGCTGACCTTAAGTCAGACAAAGCCAACTTTACATCCAGGACCCCAAGGGGCTG 120
QY 1262 GCAAGAATTCCAAAGGATGCAACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAAC 1321
DB 121 GCAAGAATTCCAAAGGATGCAACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAAC 180
QY 1322 CTGGAAGGAGGACCAAGAGAAATCAGAGCCACCTCTCTGCTGCTGGGCAAACTGT 1381
DB 181 CTGGAAGGAGGACCAAGAGAAATCAGAGCCACCTCTCTGCTGCTGGGCAAACTGT 240
QY 1382 TTTGGAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGAGAAATGTGGT 1441
DB 241 TTTGGAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGAGAAATGTGGT 300
QY 1442 GTGAGTCCACAGTAGAGATTAAAGTCCAAAGGAGTAGAATCAGCTTACAAACAGTGG 1501
DB 301 GTGAGTCCACAGTAGAGATTAAAGTCCAAAGGAGTAGAATCAGCTTACAAACAGTGG 360
QY 1502 ACCTCAACGAGGAGATGCTGCACTGAACCCACAGAGCGGAACCTCAAAAGAGAGAAA 1561
DB 361 ACCTCAACGAGGAGATGCTGCTCTCTGTAACCCACAGAGCGGAACCTCAAAAGAGAGAAA 420
QY 1562 GCAACCAAGAACCTCTCTGATGGGTTTCTCAGACAAATGTCAGTGAAGGGGATGGAG 1621

```

```

421 GCAAAACCAAGAACCTCTCTGATGGGTTTCTCAGACAAATGTCAGTGAAGGGGATGGAG 480
1622 GGATCACCCTCCTCAGAGAAATAAATGGAAAGACTCCAGTGGCCAAACATCAGACTCCA 1681
481 GGATCACCCTCCTCAGAGAAATAAATGGAAAGACTCCAGTGGCCAAACATCAGACTCCA 540
1682 CAGAAAGACTATCACCACCGCCAGAGCTGAACCAACAGGAGGACCAACAGAGGGTAAAG 1741
541 CAGAAAGACTATCACCACCGCCAGAGCTGAACCAACAGGAGGACCAACAGAGGGTAAAG 600
1742 AGGGCTCTCTGAAAGGACAAAGAGTTCAGCAGCGCCAGATGAAACAAGCAGAGAGCAACAAGC 1801
601 AGGGCTCTCTGAAAGGACAAAGAGTTCAGCAGCGCCAGATGAAACAAGCAGAGAGCAACAAGC 660
1802 AGGAAGCCAAAGAACCCAGCCAGTGCACAGAGCGGACCCAGGTGGACAGCAACTCACT-G 1860.
661 AGGAAGCCAAAGAACCCAGCCAGTGCACAGAGCGGACCCAGGTGGACAGCAACTCACTGG 720
1861 CAGATGGGGACAGCTCCAAAGAGAGACCTCAGAGCGGACCGGACAGTCCCTTGGGGGCTTC 1920
721 CAGATGGGGACAGCTCCCAAGAGACCTCAGAGCGGACCGGACAGTCCCTTGGGGGCTTC 779
1921 TTTAAAGGCTTGGGACCAAAAGCGGAT 1946
780 TCTAAAGGCTTGGGACCAAAAGCGGAT 805

```

```

RESULT 5
BX419743
LOCUS BX419743
DEFINITION BX419743 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODF020YE04 5-PRIME, mRNA sequence.
ACCESSION BX419743
VERSION BX419743.2 GI:46955113
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li.W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30765911.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 9860.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSODF020BC02QPI&c=9860.r.
FEATURES             source
   1..1089
    Location/Qualifiers
        1..1089
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="CSODF020YE04"
            /tissue_type="FETAL BRAIN"
            /dev_stage="fetal"
            /clone_lib="Homo sapiens FETAL BRAIN"
            /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
            was primed with a NotI-oligo(dT) primer. Five prime end
            enriched, double-strand cDNA was digested with Not I and
            cloned into the Not I and EcoRV sites of the pCMVSPORT 6
            vector. Library was not normalized."
ORIGIN
Query Match      24.1%; Score 737.8; DB 5; Length 1089;

```

Best Local Similarity 80.1%; Pred. No. 1.3e-109;		Matches 1013; Conservative 8; Mismatches 6; Indels 237; Gaps 4;	
Qy	751	GCAGGGATCCAC-GCTTCTCCACCTGACACAGGGGGAGCAGGAGAGCTCCCTC	809
Db	1	GCCANGATCCACAGACTTCTCCACCTGACACAGGGGGAGCAGGAGAGCTCCCTC	60
Qy	810	CAAGCCAGGACTCCAGCTTTTGTGCAAAATCTTCAAGCTGGACAGGAGCAGGAGAA	869
Db	61	CAAGCCAGGACTCCAGCTTTTGTGCAAAATCTTCAAGCTGGACAGGAGCAGGAGAA	120
Qy	870	GGTCCAGGTGACAGCCCAACAGGAAGCAAGAGGGCAGAGCATCAAGACRAGTGGATGA	929
Db	121	GGTCCAGGTGACAGCCCAACAGGAAGCAAGAGGGCAGAGCATCAAGACRAGTGGATGA	180
Qy	930	GGTCTCGGCTTATCAGGGAGTCCGATGATGTCCTGACGGGAAGGACATAGTTGACGG	989
Db	181	GGTCTCGGCTTATCAGGGAGTCCGATGATGTCCTGACGGGAAGGACATAGTTGACGG	240
Qy	990	CAAGGAAAGAGGACAGAACTTGGAACTGCGGATTTGCTCTCTCTGGGGACCCAGA	1049
Db	241	CAAGGAAAGAGGACAGAACTTGGAACTGCGGATTTGCTCTCTCTGGGGACCCAGA	300
Qy	1050	AGGACTGGAGACTGCAAGAGCAGATTCCAGGCGACAGCTATAGCAGAGATAATAATTC	1109
Db	301	AGGACTGGAGACTGCAAGAGCAGATTCCAGGCGACAGCTATAGCAGAGATAATAATTC	360
Qy	1110	CATCATGATTTCTTAAACTCTGGTTTACCTTAACAAAGCTGAACAAAAAGGACCC	1169
Db	361	CATCATGATTTCTTAAACTCTGGTTTACCTTAACAAAGCTGAACAAAAAGGACCC	420
Qy	1170	AGAAGACACGGGTGCTGAAAAGTCAACCAACCTTCAAGCTGAAGTAAAGCAAGC	1229
Db	421	AGAAGACACGGGTGCTGAAAAGTCAACCAACCTTCAAGCTGAAGTAAAGCAAGC	480
Qy	1230	CAACTTTACATCCAGGAGACCAAGGGCTGGCAAGATTCCAAAGGATCAACCCATC	1289
Db	481	CAACTTTACATCCAGGAGACCAAGGGCTGGCAAGATTCCAAAGGATCAACCCATC	540
Qy	1290	GGGGCACACAGTCCGTGACAAACCCCTGAACCTGCGAAGGAGCACCAG-GAGAAAT	1348
Db	541	GGGGCACACAGTCCGTGACAAACCCCTGAACCTGCGAAGGAGCACCAGAGAGAAAT	600
Qy	1349	CAGGACCCACCTCTCTGCTCTGGCAACCTGTTTGGAAAAGTCAGTTAAAGAGGACT	1408
Db	601	CAGGACCCACCTCTCTGCTCTGGCAACCTGTTTGGAAAAGTCAGTTAAAGAGGACT	660
Qy	1409	CAGTCCCCACAGGTGCGGAGGAGATGTGGTGTGTGAGTCAACGATAGAGATTATAAGT	1468
Db	661	CAGTCCCCACAGGTGCGGAGGAGATGTGGTGTGTGAGTCAACGATAGAGATTATAAGT	686
Qy	1469	CAAAGGAGTAGAATCAGCCTTACAAACAGTGGACCTCAACGAAGGAGATGCTGCACCTG	1528
Db	687	-----	686
Qy	1529	AACCCACAGAGCGAAACTCAAGAGAGAAAGCAACCAAGAACCTCTCTGATGGGT	1588
Db	687	-----	686
Qy	1589	TTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCTCAGAGAAATAAATG	1648
Db	687	-----	686
Qy	1649	GGAAAGACTCCAGTCCCAACATCAGACTCCACAGAAAGACTATCAACGGCCAGAGC	1708
Db	687	-----ACATCAGACTCCACAGAAAGACTATCAACGGCCAGAGC	726
Qy	1709	CTGACCAACAGGAGCACACAGAGGTTAAGAGGGCTCTCCGAAAGGACAGAGTCAAG	1768
Db	727	CTGACCAACAGGAGCACACAGAGGTTAAGAGGGCTCTCCGAAAGGACAGAGTCAAG	786
Qy	1769	CAGCCGAGATGACAAAGCAGAGAGCAACAGAGGAGGCCAAAGAACCCAGCCCAAGTCA	1828

Db	787	CAGCCGAGATGAACAAGCAGAGAGCAACAAGCAGGAAGCCAAAGAACCCAGCCAGTGCA	846
Qy	1829	CAGAGAGGCCACGGTGGACACGAACCTCACTGCAGAAATGGGACCAAGCTCCAAAAGAGAC	1888
Db	847	CAGAGAGGCCACGGTGGACACGAACCTCACTGCAGAAATGGGACCAAGCTCCAAAAGAGAC	906
Qy	1889	CTGAGAGAGGCCACGGTGGGCGCTTCTTT-AAAGGCTTGGGACCAAAAGCGGATG	1947
Db	907	CTGAGAGAGGCCACGGTGGGCGCTTCTTTAAAGGCTTGGGACCAAAAGCGGATG	966
Qy	1948	TTGATGCTCAAGTGCACAGAGCCAGATTCATCCGACAGCTTGGCAAAACCAAGTAA	2007
Db	967	TTGATGCTCAAGTGCACAGAGCCAGATTCATCCGACAGCTTGGCAAAATCCAAARWAA	1026
Qy	2008	ACAA 2011	
Db	1027	AMAA 1030	
RESULT 6			
CN646528			
LOCUS			
DEFINITION			
ILLUMIGEN_MQO 26475 Katze_MBR Macaca mulatta cDNA clone IBIUW.8476			
5', similar to Bases 146 to 797 highly similar to human BCAS1			
(Hs.400556), mRNA sequence.			
ACCESSION			
CN646528			
VERSION			
CN646528.1 GI:47159971			
KEYWORDS			
EST.			
SOURCE			
Macaca mulatta (rhesus monkey)			
ORGANISM			
Macaca mulatta			
REFERENCE			
1 (bases 1 to 897)			
Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magnus,C.L.			
Large-scale Rhesus Macaque cDNA Sequencing			
Unpublished (2003)			
JOURNAL			
COMMENT			
Contact: C. Magnus			
Illumigen Biosciences Inc.			
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA			
Tel: 2063780400			
Fax: 2063780408			
Email: cmagness@illumigen.com			
Sequenced on 2004.03.10. 746 Q20 bases. Contains sequence similar			
to Alu			
PCR Primers			
FORWARD: CCTCCTAAAGGGAACAAA			
BACKWARD: CACTATAGGCGAATGGGTA			
Insert Length: 897 Std Error: 0.00			
Plate: CL000144 row: G column: 09			
Seq primer: CCTCCTAAAGGGAACAAA			
POLYA-No.			
Location/Qualifiers			
1..897			
/organism="Macaca mulatta"			
/mol_type="mRNA"			
/strain="Indian"			
/db_xref="taxon:9544"			
/clone="IBIUW.8476"			
/sex="female"			
/dev stage="adult"			
/lab host="E. coli SOLR"			
/clone.lib="katze_MBR"			
/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: Ecor I;			
Site 2: Xho I; Created from Stratagene ZAP-cDNA Synthesis			
kit (catalog #200400) and ZAP-CDNA Gigapack III Gold			
Cloning Kit (Catalog #200450)"			
ORIGIN			
Query Match 22.7%; Score 696.4; DB 7; Length 897;			
Best Local Similarity 95.2%; Pred. No. 6.7e-103;			
Matches 718; Conservative 0; Mismatches 36; Indels 0; Gaps 0;			

SEGI PRSBSNVKSSCOTSNSVEKTPSPPEPAGTAQKNKETSSSKDKSVDKKS
TENSOKNGKQREBPAPCVOPPTVEANAMTGDKTPKSEKRQSLGFLKGLGPKR
MSDAQVOTDPSVIGPVGKSK
2317..2322
/note="putative"
2339
/note="putative"

polyA_signal

polyA_site

ORIGIN

Query Match 19.4%; Score 595.6; DB 3; Length 2339;

Best Local Similarity 64.8%; Pred. No. 1.le-86;

Matches 1095; Conservative 0; Mismatches 414; Indels 181; Gaps 7;

Qy	796	GGAGAAAGCTCCCTCCCAAGCCCAAGGACTCCAGCTTTTGGACAAATCTTCAAGCTGGAC	855
Db	55	GCAGAAAGCCCGCCGACGACAAAGGATTTGCGCTTTTGAACAGATTTCTTAAACTGGAC	114
Qy	856	AAGGACAGGAAAGGTGCCAGGTGACAGCCAAAC---AGGAAGCCCAAGAGGCGCAGAGCAT	912
Db	115	AAAGGAAGAAAGCGCCCGGTGAACAGCCAGCCCAAGAAAGCGAAAGGCTCGGAAGAC	174
Qy	913	CAAGACAAGGTGGATAGAGTTCCTGGCTTTATCAGGGCAGTCCGATGATGTCCTGCGAGG	972
Db	175	CCAGAACAGGCCACAGAGGCTCCTGCGTGCAGGGAATCCCATGTTGTTCTCTGCGAGG	234
Qy	973	AAGGACATAGTTGACGGCAAGGAAAGAAAGAGGACAAAGTGGAACTTCGCGGATGCTCT	1032
Db	235	GAGGACATAGTCGACAGCGAGCAGAA---GAGGACAAAGCGTTGACACTCTGAGTTATTCT	291
Qy	1033	GTCCCTGGGACCCCAAGAGGACTGGAGACTGCAAGAGGACGATTCGCCAGCGCAGCAGCTATA	1092
Db	292	GTCCCTGGGATCCAGAGTGTCCGGGACACGAAAGGAGGACCCGCGAGTGTGATACC	351
Qy	1093	GCAGAGATAATATTCATCATGATGTTCTTTAAACCTCTGGTTTACCTAAACAAGCT	1152
Db	352	ACAGAGACAGCAGCTCCATCATGAGCTTCTTCAAGACACTGGTTTACCTTAAACAAGCT	411
Qy	1153	GAACAAAGGACCCAGAGACAC	1178
Db	412	GAACAAAGAGGACCCAGAGACACAGAAAGGCAACCAAGGACAGAGTGTCTGTGATGA	471
Qy	1179	-----	1178
Db	472	CACGCTGTCGCCAAGATGTCGAGAGCAGGCTAAAGCAAGAAAGGCGCTGGAC	531
Qy	1179	-----GGGTGCTGAAAAGTCA	1194
Db	532	AGCCCCAGGCTAGGACTCTCTTTTAGAAGCTCTTTAGGCACAAAGGATACTAGAACTCA	591
Qy	1195	CCACCACTTCAGCTGACCTTAAGTCAGACAAAGCCAACTTTACATCCAGGAGACCCAA	1254
Db	592	CCACGACTTCAGCCAAATCTCAAGTCGGAACAAAGCCAACTTCAACCCCGAGAGACCCGA	651
Qy	1255	GGGGTGGCAAGAAATTCACAAAGGATGCAACCATCGGGGACACACAGCTCGTGACACC	1314
Db	652	GGGAAGACAAAGGCTACCAAAAGCTGACAGCCCCCACCAGCCACTC-----CTGAACCC	705
Qy	1315	CCTGAACCTCGGAAGGACCAAGGAGAGAAATCAGGACCCACCTCTCTGCTCTGGGC	1374
Db	706	ACCAGCAAGGAGAGACAGCGGCAAGGAGAGGCGAGGCGCCACCTCACTACCGCTGGGA	765
Qy	1375	AAACTGTTTTGGAAGAGTCAAGTTAAAGAGACTCAGTCCCGACAGGTGCGGAGAGAAAT	1434
Db	766	AAGTTGTTTTGGAAGAGTCAAGTTAAAGAGGATACACTTCCACAGGTGTCAGAGAGAAC	825
Qy	1435	GTGTTGTCGTAGTACCAAGTACAGATTATAAGTCCRAAGGAGTAGAATCAGCTTACAA	1494
Db	826	CGGTTGTCGTAGTACCAAGTACAGACCGTAGGCTTTGAGGAGTAGAATCAGCTTACAA	885
Qy	1495	ACAGTGGACCTCAACGAAGGAGAGTCTGCACCTGAACCCACAGAGGCGAAACTCAAGA	1554
Db	886	ACTGTGATCTCAGTGA---AGAGACCCAGCTGACCCACAGAGCTAAAGTCAAGAA	942

Qy	1555	GAAGAAAGCAAAACCAAGAACCTCTCTGATGGCTTTCTCAGACAAATGTCTAGTGAAGGG	1614
Db	943	GAAGCAAAACCCCGGAAGACCCCTCTGATGGCTTTCTCAGACAAATGTCTAGTGAAGCG	1002
Qy	1615	GATGAGGAGTACCCACTCAGAAAGAAATAAATGGAAAGACTCCAGCTGCCAAACATCA	1674
Db	1003	AGCGAAGGATCCCGCTCGGAAGAAAGTAACTGTAAGACTCCAGCTGCCAAACGTC	1062
Qy	1675	GACTCCACAGAAAGACTATCAGCCGACAGGCTTGAAACCAAGGAGACACACAGAAAG	1734
Db	1063	AACCTCGTGGAAAGACGCCCTCAGCCCGACAGGCGCGCGGAAACAGCTCAGAAA	1122
Qy	1735	GGTAAAGAGGGCTCC-----TCGAAGGACAAAGCTTCAGCACCCGAG	1776
Db	1123	AACAGGAGAGCTCTCTCTCGAAGGACAAAGAAATCAGTGGACAAAGTCTAGCGACTGAG	1182
Qy	1777	ATGAACAGCAGAGAGCAACAGCAGGAAGCAAAAGAACCCAGCCAGTGCACACAGAG	1836
Db	1183	AACAGCAGCAGAGAGACGGCAACAGAGAGTCAAGAGACCGCGCGCTGTGTGACGCG	1242
Qy	1837	GCCACGTTGGACACGAATCTCAGAAATGGGACAAAGCTCCAAAGAGACCTCGAAG	1896
Db	1243	CCACAGTGGAGGCAACGCAATGCAGACTGGGACAAAGACCCCAAGAAAGTCTGAGAAG	1302
Qy	1897	CGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTTGGACCAAGACCGGATGTTGATGCT	1956
Db	1303	CGGCGACAGTCCCTCGGGGGCTTCTGAAAGGGCTTGGACCAAGCGGATGTCGATGCT	1362
Qy	1957	CAAGTGCAACAGACCCAGATATCCATCGGACAGTTGGCAAAACCAAGTAAACAAATCAG	2016
Db	1363	CAGTGCACAAAGGACCCGCTCTCATCGGACAGTTGGAAATCCAGTAAACAAATCA	1422
Qy	2017	CACGGTCCCAACAGGTTCTCTGCGCAACAGATGTTGTTCTCTTCTTCTCTCTCTCT	2076
Db	1423	CCAGTTCCTCACTGGGTTCTCTCTCCAC-----TCCATCTCTCTCTCTCTCTCTCT	1472
Qy	2077	CACACAGCTCCATGATATATCTCTGATGGCCCAAGTAAATGAAATCTCGCTAGAAAT	2136
Db	1473	ACTGTATATATATTTCT	1532
Qy	2137	TAAGCCCGAGCTGTTGATATTTAGAGTGTATTTATTTAGCTCTCTGCTCCAGTCTTTCTG	2196
Db	1533	TCACTGAAATTTGTTAGATCGAGGTTATTTATTTAGCTCTCTGCTCCAGTCTTTCTCTG	1592
Qy	2197	GCAATTAACAGTAAGATGTTTAGCAGGTCACTTGTGGGTGAGAGAGTTCAGATGATC	2256
Db	1593	GCAATTAAGTGAATAAATGGCTTAGCAGGTTGACCAAGTTGATCCAAAGGAGCTGGTGAGT	1652
Qy	2257	ACCAAGCAGGAAAGGAGGAGGAAATAGAGAAATGTTTGGGTTAAGTGAATGAAATGGCAG	2316
Db	1653	GCCAGGTGGGAGGAGGAGTAGGTAGTAAAGGCATTTGTGTTAAGTTATGAATAATGGC	1712
Qy	2317	TGGTGGCCGG 2326	
Db	1713	AGGTATCCAG 1722	

RESULT 9
LOCUS AL120048 610 bp mRNA linear EST 04-SEP-2003
DEFINITION DKEP2761K142_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
ACCESSION DKEP2761K142_5', mRNA sequence.
VERSION AL120048
KEYWORDS AL120048.1 GI:5925947
SOURCE EST.
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 610)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE EST (Ottenwaelder, et al.)

AL120048 610 bp mRNA linear EST 04-SEP-2003
DKEP2761K142_r1 761 (synonym: hamy2) Homo sapiens cDNA clone
DKEP2761K142_5', mRNA sequence.
AL120048.1 GI:5925947
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 610)
Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Ottenwaelder, et al.)

JOURNAL
COMMENT

Unpublished (1999)

Contact: MIPS

MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer

Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by MediGenomix (Martinsried/Germany) within the cDNA

sequencing consortium of the German Genome Project. No s1 sequence

available.

This clone (DKFZp761K142) is available at the RZPD in Berlin.

Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059

Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source

1. 610

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="DKFZp761K142"

/tissue type="amygdala"

/dev stage="adult"

/lab_host="DH10B"

/clone_lib="761 (synonym: hamy2)"

/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.4%; Score 594.8; DB 1; Length 610;

Best Local Similarity 99.5%; Pred. No. 2e-86;

Matches 607; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Qy 1719 AGGAGCACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGTGCAGCGCGAGAT 1778

Db 1 AGGAGCACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGTGCAGCGCGAGAT 60

Qy 1779 GAACAGCAGAAAGACACAGCAGGAGCCAAAGAACCCAGCCCGAGTGCACAGAGCAGGC 1838

Db 61 GAACAGCAGAAAGACACAGCAGGAGCCAAAGAACCCAGCCCGAGTGCACAGAGCAGGC 120

Qy 1839 CACGGTGACACAGCACTCAGTGCAGATGGGACAAAGTCCAAAAGAGACCTGAGAAGCG 1898

Db 121 CACGGTGACACAGCACTCAGTGCAGATGGGACAAAGTCCAAAAGAGACCTGAGAAGCG 180

Qy 1899 CGACAGTCCCTTGGGGCTCTTTAAAGGCGCTGGACCAAAGCGGATGTTGGATGCTCA 1958

Db 181 CGACAGTCCCTTGGGGCTCTTTAAAGGCGCTGGACCAAAGCGGATGTTGGATGCTCA 240

Qy 1959 AGTGCAACACAGACCCAGTATCATCGGACCCAGTTGGCAAAACCCAAAGTAAACATCAGCA 2018

Db 241 AGTGCAACACAGACCCAGTATCATCGGACCCAGTTGGCAAAACCCAAAGTAAACATCAGCA 300

Qy 2019 CGGTTCCCAACAGGTTCTCTGCGCACCAAGATGTTCTCTTACTCATCTCTCTCCCA 2078

Db 301 CGGTTCCCAACAGGTTCTCTGCGCACCAAGATGTTCTCTTACTCATCTCTCTCCCA 360

Qy 2079 AACAGCTCCATGATATATTTCTTGATGCCAGCAAAATGAATTTGCTTAGAAATTA 2138

Db 361 AACAGCTCCATGATATATTTCTTGATGCCAGCAAAATGAATTTGCTTAGAAATTA 420

Qy 2139 AGCCGAGCTGTTGATATTAGGTGTTATTTACGTCTCTGTCAGTCTTTCTGCGC 2198

Db 421 AGCCGAGCTGTTGATATTAGGTGTTATTTACGTCTCTGTCAGTCTTTCTGCGC 480

Qy 2199 AAATAACAGTAAAGATGTTTAGCAGGTCACCTAGTTGGGTGAGAAGAGTGCATGATCAC 2258

Db 481 AAATAACAGTAAAGATGTTTAGCAGGTCACCTAGTTGGGTGAGAAGAGTGCATGATCAC 540

Qy 2259 CAAGCAGGAAAGGGAGGAATAG -AGGAATGTGTTTCGGGTTTAAGTGTGATAAATGGCAGT 2317

Db 541 CAAGCAGGAAAGGGAGGAATAGAGGAATGTGTCGGGTTTAAGTGTGATAAATGGCAGT 600

Qy 2318 GGTGCGCGG 2327

Db 601 GGTGCGCGG 610

RESULT 10

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

C0724689 820 bp mRNA linear EST 27-JUL-2004
ILLUMIGEN MQ0_8605 Katze_MBR Macaca mulatta cDNA clone IBIUW:24007
5' similar to Bases 70 to 720 highly similar to human BCAS1
(Hs.400556), mRNA sequence.

C0724689
C0724689.1 GI:50704602
EST.

Macaca mulatta (rhesus monkey)
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
Cercopithecinae; Macaca.

1 (Bases 1 to 820)
Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
Large-scale Rhesus Macaque cDNA Sequencing
Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.

2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com

Sequenced on 2004.01.16. 530 Q20 bases. Contains sequence similar
to Alu Library Preparation: Prof. Michael Katze Lab at University
of Washington DNA Sequencing: Illumigen Biosciences Inc. For
further information, see http://www.macaque.org

PCR Primers
FORWARD: CCCTCACTAAAGGGAACAAAA
BACKWARD: CACTATAGCGGAATTGGGTA

Insert Length: 820 Std Error: 0.00

Plate: CL000089 row: G column: 09

Seq primer: CCCTCACTAAAGGGAACAAAA

POLYA=No.

Location/Qualifiers

1. 820

/organism="Macaca mulatta"

/mol_type="mRNA"

/strain="Indian"

/db_xref="taxon:9544"

/clone="IBIUW:24007"

/sex="female"

/dev stage="adult"

/lab_host="E. coli SOLR"

/clone_lib="Katze_MBR"

/note="Organ: brain; Vector: Uni-ZAP XR; Site 1: EcoR I;

Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis

kit (catalog #200400) and ZAP-CDNA Gigapack III Gold

Cloning Kit (Catalog #200450)"

ORIGIN

Query Match 19.3%; Score 592.6; DB 7; Length 820;

Best Local Similarity 93.3%; Pred. No. 4.3e-86;

Matches 652; Conservative 0; Mismatches 44; Indels 3; Gaps 3;

Qy 1667 AAACATCAGATCCACAGAAAGACTATCACACCGCAGAGCCTGAACCAACAGGAGCAC 1726

Db 68 ACACATCATATCTCGCTCAAAGAGACTATCCACACCCCTTACCTAGAGAAACAGGAGCGC 127

Qy 1727 CACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGAGTTCAGAGCCCGAGATGAACAAGC 1786

Db 128 CCCAGAGGGCAAGAGGGCTCTTCGAGGACAGAGTTCAGAGTTCAGATGACAAAC 187

Qy 1787 AGAAGAGCAACAGCAGGAAGCCAAAGAAACAGGCCAGTGCACAGAGCAGCCAGCTGG 1846

Db 188 AGAAGAGCAACAGCAGGAAGCCAAAGAAACAGGCCAGTGCACAGAGCAGCCAGCTGG 247

Qy 1847 ACACGAATCTCTCGAGTGGGACAGCTCCAAAAGAGACCTCAGAGCGGCGAGCAGT 1906

Db 248 ACACGAATCTCTCGAGTGGGACAGCTCCAAAAGAGACCTCAGAGCGGCGAGCAGT 307

Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..598
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="ADR06635"
/tissue type="brain"
/clone_lib="Sugano cDNA library, brain"

ORIGIN
Query Match 19.0%; Score 583.4; DB 5; Length 598;
Best Local Similarity 99.3%; Pred. No. 1.4e-84;
Matches 595; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy 670 CTTCCGGTGGCAGCTGGACCGGGGAGGACACAGATAAAACCCAGGGGACGCCCCGGCC 729
Db 1 CTTCCGGTGGCAGCTGGACCGGGGAGGACACAGATAAAACCCCA-GGCACGCCCCGGCC 59
Qy 730 CAAGACAGGTCCTCTCTGCCCGCAGGATCCCGCTTCCACCTGACACAGGGGA 789
Db 60 CAAGACAGGTCCTCTCTGCCCGCAGGATCCCGCTTCCACCTGACACAGGGGA 119
Qy 790 GCAGAGAGAAAGCTCCCTCCAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAG 849
Db 120 GCAGAGAGAAAGCTCCCTCCAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAG 179
Qy 850 CTGCAAGAGGACAGGAAAGGTGCCAGTGAAGCCCAAGGAGGAGGAGGAG 909
Db 180 CTGCAAGAGGACAGGAAAGGTGCCAGTGAAGCCCAAGGAGGAGGAGGAG 239
Qy 910 CATCAAGACAGGTGGATGAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGCCGTGCA 969
Db 240 CATCAAGACAGGTGGATGAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGCCGTGCA 299
Qy 970 GGGAGGACATAGTTGACGGCAAGGAAAGAGGACAGAACTTGGAACTGCGGATTGC 1029
Db 300 GGGAGGACATAGTTGACGGCAAGGAAAGAGGACAGAACTTGGAACTGCGGATTGC 359
Qy 1030 TCTGTCTCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCCAGGCGAGCAGCT 1089
Db 360 TCTGTCTCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCCAGGCGAGCAGCT 419
Qy 1090 ATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGGTTTCACTTAACAAA 1149
Db 420 ATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGGTTTCACTTAACAAA 479
Qy 1150 GCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCACTTCAGCT 1209
Db 480 GCTGAAACAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCACTTCAGCT 539
Qy 1210 GACCTTAAGTCAGACAAAGCCAACTTTATCTCCAGGAGACCAAGGGGCTGGCAAGAA 1268
Db 540 GACCTTAAGTCAGACAAAGCCAACTTTATCTCCAGGAGACCAAGGNGCTGGNCAGAA 598

RESULT 13
BP342407
LOCUS
DEFINITION BP342407 Sugano cDNA library, stomach mucosa Homo sapiens cDNA
ACCESSION BP342407
VERSION BP342407.1 GI:52272209
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
Suzuki, Y., Yamashita, R., Shirota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yezuki@ims.u-tokyo.ac.jp.

FEATURES
source
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="STM04831"
/tissue type="stomach mucosa"
/clone_lib="Sugano cDNA library, stomach mucosa"

ORIGIN
Query Match 19.0%; Score 581.4; DB 5; Length 583;
Best Local Similarity 99.8%; Pred. No. 3e-84;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 146 AGCCAGGACTGGGACGACGCACTGGAGACCCAGGACCCCTGTGCAGGAGCAGCTCCGGG 205
Db 1 AGCCAGGACTGGGACGACGCACTGGAGACCCAGGACCCCTGTGCAGGAGCAGCTCCGGG 60
Qy 206 TGACACAGGAGGACTGAAGATACTCCACAGGGGCTCAGCAGGAGCAATGGTAAACAAA 265
Db 61 TGACACAGGAGGACTGAAGATACTCCACAGGGGCTCAGCAGGAGCAATGGTAAACAAA 120
Qy 266 TGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGAGAGACTTACCAAG 325
Db 121 TGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAATGAACAGAGAGAGACTTACCAAG 180
Qy 326 ACACAGGCTCTGCTCAACGGGGTTCAGTGGTGTGCGACCCACACAGCTTCAGCACT 385
Db 181 ACACAGGCTCTGCTCAACGGGGTTCAGTGGTGTGCGACCCACACAGCTTCAGCACT 240
Qy 386 TAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGCG 445
Db 241 TAGAGGAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGCG 300
Qy 446 AGACACAGGAGTAAGTGTGTTGGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCA 505
Db 301 AGACACAGGAGTAAGTGTGTTGGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCA 360
Qy 506 AACCCGGGACACAGCTGCTAAATCTGTTTTTTTCTTGATGCTCTCTCGGCTGTACCAG 565
Db 361 AACCCGGGACACAGCTGCTAAATCTGTTTTTTTCTTGATGCTCTCTCGGCTGTACCAG 420
Qy 566 GACCTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTACGT 625
Db 421 GACCTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTACGT 480
Qy 626 CCAATAAAGTCTCAGCGCAACAAAGACCAAGTGAAGCTGGACACTTCCCGTGGCAGCTG 685
Db 481 CCAATAAAGTCTCAGCGCAACAAAGACCAAGTGAAGCTGGACACTTCCCGTGGCAGCTG 540
Qy 686 GACCCGGGACACACAGATAAAACCCAGGACGCCCCCGGC 728
Db 541 GACCCGGGACACACAGATAAAACCCAGGACGCCCCCGGC 583

RESULT 14
BM702087
LOCUS
DEFINITION UI-E-CQ1-aey-i-21-0-UI.r1 UI-E-CQ1 Homo sapiens cDNA clone

UI-E-CQ1-aey-i-21-0-UI 5', mRNA sequence.

ACCESSION BM702087
 VERSION BM702087.1 GI:19015345
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 589)
 REFERENCE Ronald,M.F., Lennon,G. and Soares,M.B.
 AUTHORS Normalization and subtraction: two approaches to facilitate gene
 TITLE discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@uiowa.edu
 Tissue Procurement: Dr. Gregg Hageman
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 The following repetitive elements were found in this cDNA
 sequence: 390-589, >ALU
 Seq primer: M13 Reverse

FEATURES

source
 Location/Qualifiers
 1..589
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="UI-E-CQ1-aey-i-21-0-UI"
 /tissue_type="optic nerve"
 /dev_stages="adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-E-CQ1"
 /note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
 modified polylinker; Site 1: EcoR I; Site 2: Not I;
 UI-E-CQ1 is a normalized cDNA library containing the
 following tissue(s): optic nerve. The library was
 constructed according to Bonaldo, Lennon and Soares,
 Genome Research, 6:791-806, 1996. First strand cDNA
 synthesis was primed with an oligo-dT primer containing a
 Not I site. Double stranded cDNA was ligated to an EcoR I
 adaptor, digested with Not I, and cloned directionally
 into pT7T3-Pac vector. The oligonucleotide used to prime
 the synthesis of first-strand cDNA contains a library tag
 sequence that is located between the Not I site and the
 (dT)18 tail. The sequence tag for this library is
 CCATTAAGTG. This library was created for the program, Gene
 Discovery in the Visual System, supported by National Eye
 Institute (NEI)."

ORIGIN

Query Match 18.9%; Score 581; DB 4; Length 589;
 Best Local Similarity 99.2%; Pred. No. 3.5e-84;
 Matches 584; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 1932 GGGACCAAGCGGATGTTGGATGCTCAAGTGCAGAACAGACCCAGATTCATCGGACCAAGT 1991
 DB 1 GGGACCAAGCGGATGTTGGATGCTCAAGTGCAGAACAGACCCAGATTCATCGGACCAAGT 60
 QY 1992 TGGCAAAACCAAGTAAACAATACAGACGGTTCACCAAGGTTCTTCGCCACCAAGATG 2051
 DB 61 TGGCAAAACCAAGTAAACAATACAGACGGTTCACCAAGGTTCTTCGCCACCAAGATG 120
 QY 2052 TGTTCCTCTTACTCTCTCTCCCAACAGCGTCCATGTATATATTCTTCTGTATGCC 2111

Db 121 TGTTCCTCTTACTCTCTCTCCCAACAGCGTCCATGTATATATATTCTTCTGTATGCC 180
 QY 2112 AGCAAAATGAAATTTGCGCTAGAAAATTAAGCCGAGCTGTTGATATTAGGTTGATTATT 2171
 Db 181 AGCAAAATGAAATTTGCGCTAGAAAATTAAGCCGAGCTGTTGATATTAGGTTGATTATT 240
 QY 2172 TACGTCCTCTGGTCCAGTCTTTTCTGGCAAAATTAACAGTAAGATGCTTTAGCAGGTCACCT 2231
 Db 241 TACGTCCTCTGGTCCAGTCTTTTCTGGCAAAATTAACAGTAAGATGCTTTAGCAGGTCACCT 300
 QY 2232 AGTTGGGTTCAGAAAGATGCGATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTGT 2291
 Db 301 AGTTGGGTTCAGAAAGATGCGATGATCACCAGCAGGAAAGGAGGGAATAGAGGAATGTGT 360
 QY 2292 TCGGGTTAAGTGAATAAATGCGAGTGTGCGCGGCGTGGTGGCTTCGCTCGCTGTAATCT 2351
 Db 361 TCGGGTTAAGTGAATAAATGCGAGTGTGCGCGGCGTGGTGGCTTCGCTCGCTGTAATCT 420
 QY 2352 CAGCACATTTGGGAGCGCGAGGCGAGTGTGATCACCTGAGTTCAGGAGTTCAGACTAGCT 2411
 Db 421 CAGCACATTTGGGAGCGCGAGGCGAGTGTGATCACCTGAGTTCAGGAGTTCAGACTAGCT 480
 QY 2412 GGCCAAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGCTGGC 2471
 Db 481 GGCCAAACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGCTGGC 540
 QY 2472 ACACACCTGTAGTCCAGCTACTCGGAGCGCAACGACGAGAACCGCT 2520
 Db 541 ACACACCTGTAGTCCAGCTACTCGGAGCGCAACGACGAGAACCGCT 589
 RESULT 15
 LOCUS BX486692
 DEFINITION DKFZp686P08252_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 ACCESSION BX486692
 VERSION BX486692.1 GI:31950607
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 583)
 AUTHORS Bahr,A., Lauber,J., Mewes,H.W., Weil,B., Amid,C., Osanger,A.,
 Fobo,G., Han,M. and Wiemann,S.
 EST (Bahr,A., Lauber,J., Mewes,H.W., Weil,B., et al.)
 JOURNAL Unpublished (2003)
 COMMENT Contact: MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
 consortium of the German Genome Project.
 No s1 sequence available.
 This clone (DKFZp686P08252) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
 Location/Qualifiers
 1..583
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZp686P08252"
 /dev_stages="adult"
 /lab_host="DH10B"
 /note="Vector: pTriplex2; (synonym: hlcc3)"
 /note="Vector: pTriplex2; (synonym: hlcc3)"
 cDNA-collection"

ORIGIN

Query Match		18.7%;	Score 573.2;	DB 5;	Length 583;
Best Local Similarity		99.5%;	Pred. No. 6.5e-83;		
Matches 575;		Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	127	GGCAGGAGCTCCACCGCAGCCAGGCACTGGGAGCAGCGCACTGGAGACCCAGGACCCCT	186		
Db	6	GGAGCCAGGCTCCACCGCAGCCAGGCACTGGGAGCAGCGCACTGGAGACCCAGGACCCCT	65		
Qy	187	GTGCAGGAGCAGCTCCCGGTGACAGGAGGACTGAAGATACTCCACAGGGGTTCAGCA	246		
Db	66	GTGCAGGAGCAGCTCCCGGTGACAGGAGGACTGAAGATACTCCACAGGGGTTCAGCA	125		
Qy	247	GGAGCAATGGGTAAACCAATAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAATCAACCA	306		
Db	126	GGAGCAATGGGTAAACCAATAGTGTTCGCCAAAGAGTTGAAGACCAAGAGAATGAACCA	185		
Qy	307	GAAGCAGAGACTTACCAGGACAAACGCTCTGCTCTGAACGGGGTTCAGTGGTGGTGTGCG	366		
Db	186	GAAGCAGAGACTTACCAGGACAAACGCTCTGCTCTGAACGGGGTTCAGTGGTGGTGTGCG	245		
Qy	367	ACCCACACAGTTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAAT	426		
Db	246	ACCCACACAGTTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGACGGATAAT	305		
Qy	427	GTGGCCACTTCTTCCCCGAGACAAACGAGATAAGTGTGTGGGATGCCAAACGGAAG	486		
Db	306	GTGGCCACTTCTTCCCCGAGACAAACGAGATAAGTGTGTGGGATGCCAAACGGAAG	365		
Qy	487	AATCTTGGGAAGAGGCCAAACCCGAGGACCGACCTGCTAAATCTCGTTTTTCTTGTATG	546		
Db	366	AATCTTGGGAAGAGGCCAAACCCGAGGACCGACCTGCTAAATCTCGTTTTTCTTGTATG	425		
Qy	547	CTCTCTCGGCTGTACACGAGGAGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCA	606		
Db	426	CTCTCTCGGCTGTACACGAGGAGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCA	485		
Qy	607	GTGAAGCTTGATGTTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGG	666		
Db	486	GTGAAGCTTGATGTTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCCAAGTGAGAGCTGG	545		
Qy	667	ACACTTCGGTGGCAGCTGGACCGGGGAGGACACAGA	704		
Db	546	ACACTTCGGTGGCAGCTGGACCGGGGAGGACACAGA	583		

Search completed: July 3, 2005, 16:40:15
Job time : 9707.27 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.
GENCORE VERSION 3.1.1.6

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 5109.25 Seconds
(without alignments)
11428.001 Million cell upd.

Title: US-08-731-499-7

Perfect score: 1205

Sequence: 1 GCGCGGTGAGTCGCCCCC.....TCAGTGTTTGGTTAAGTGG 1205

Scoring table;. IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Minimum DB seq	length: 0
Maximum DB seq	length: 2000000000

Post-processing: Minimum Match 0%

Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: a'b ba:

```
1: gb_ba: *
2: ab_bt a: *
```

2: gb_neg:*

3: gb_III: *

4: gb_om: ★
5: gb_ov: ★

6: ab_pat: *

7: qb ph: *

8: gb pl: *

9: gb_pr: *

10:- gb_ro:*

11: gb_sts:

12: gb_sy:*

13: gb_un:*

```
14: gb_vi:*
```

•

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1201	99.7	1205	6	AR070330	Sequence
2	1201	99.7	1205	6	BD085731	Genes fro
C 3	844.6	70.1	122764	9	HSU031D17	AL109984 Homo DNA
C 4	742.8	61.6	173718	9	AC073575	AC073575 Homo sapi
5	739.8	61.4	892	9	HSER28	X94910 Homo sapien
6	738.6	61.3	889	6	C0728979	C0728979 Sequence
7	635.6	52.7	786	9	CR541667	CR541667 Homo sapi
8	545	45.2	1258	10	BC017125	BC017125 Mus muscu
9	496.6	41.2	1139	6	AX827789	AX827789 Sequence
10	496.6	41.2	1139	10	RNU36482	U10264 Rattus norv
11	479.4	39.8	1084	10	RNERP29PR	U10264 R.norvegicu
12	404.8	33.6	438	6	C0692426	C0692426 Sequence
13	403	33.4	464	11	G25553	G25553 human STS E
14	403	33.4	464	11	G27410	G27410 human STS S
15	402.2	33.4	500	6	AX687897	AX687897 Sequence
16	402.2	33.4	500	6	BD027507	BD027507 Sequence
C 17	358.8	29.8	140936	10	AC129215	AC129215 Mus muscu
18	295.8	24.5	9664	10	AY004254	AY004254 Rattus no
C 19	295.8	24.5	216476	2	AC095682	AC095682 Rattus no

ALIGNMENTS

[illegible]

Db 241 TTCCAAAGCAAGTTCGCTCGTGAAGTTCCGACACCAGTACCCTACGGTGGAAGCA 300
Qy 301 GGATGAGTTCAAGGCTCTTCTGAAACCTCGGCTCCAGCGATGATCTCTTGTCGTCAGAG 360
Db 301 GGATGAGTTCAAGGCTCTTCTGAAACCTCGGCTCCAGCGATGATCTCTTGTCGTCAGAG 360
Qy 361 GTGGGATCTCAGATTTATGTGACAACTGAACTGAGCTGAGTGAAGAAATACAACTGG 420
Db 361 GTGGGATCTCAGATTTATGTGACAACTGAACTGAGCTGAGTGAAGAAATACAACTGG 420
Qy 421 ACAAAGAGAGCTACCCATCTTCTAGCTCTTCCGGGATGGGACTTTGAGAACCCAGTCCC 480
Db 421 ACAAAGAGAGCTACCCATCTTCTAGCTCTTCCGGGATGGGACTTTGAGAACCCAGTCCC 480
Qy 481 ATACACTGGGCGAGTTAGTTGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGGTCTTAC 540
Db 481 ATACACTGGGCGAGTTAGTTGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGGTCTTAC 540
Qy 541 CTAGGTATGCTGTGCTGCTGTATACGACGCTGCGCGGGGAGTTTCATCAGGGCCT 600
Db 541 CTAGGTATGCTGTGCTGCTGTATACGACGCTGCGCGGGGAGTTTCATCAGGGCCT 600
Qy 601 CTGCTGTGGAGCGCCAGGCTCTTGAAGAGGGGCAAGATACCTCTCAAGTGTGAA 660
Db 601 CTGCTGTGGAGCGCCAGGCTCTTGAAGAGGGGCAAGATACCTCTCAAGTGTGAA 660
Qy 661 GGAGACTCAGAGAGTGGCGCGCAATACCTGAGATCATGGGAGATCTTAGACCA 720
Db 661 GGAGACTCAGAGAGTGGCGCGCAATACCTGAGATCATGGGAGATCTTAGACCA 720
Qy 721 GGGAGACTTCCAGCATCAGAGATGACACGATCGCCAGGCTGATTGAGAAGAACAGA 780
Db 721 GGGAGACTTCCAGCATCAGAGATGACACGATCGCCAGGCTGATTGAGAAGAACAGA 780
Qy 781 TGAGTGACGGCAGAGAGAGTCCAGAGAGCTTAAACATCTGAGCTGCTTCCAGAA 840
Db 781 TGAGTGACGGCAGAGAGAGTCCAGAGAGCTTAAACATCTGAGCTGCTTCCAGAA 840
Qy 841 GAAGGGGCGCAGAAAGAGGAGCTGTAAGAGGCTGCTGTGATTTCCAGGGTTGGTG 900
Db 841 GAAGGGGCGCAGAAAGAGGAGCTGTAAGAGGCTGCTGTGATTTCCAGGGTTGGTG 900
Qy 901 GGGGTAGGGAGGGANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGY 960
Db 901 GGGGTAGGGAGGGANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGY 960
Qy 961 MSKGGGAAAAGGGTACTAACCCAGATTTGAGCCCTGAGTATGCCCTGGACATTTGATGC 1020
Db 961 MSKGGGAAAAGGGTACTAACCCAGATTTGAGCCCTGAGTATGCCCTGGACATTTGATGC 1020
Qy 1021 TAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGATAGCTGGGACATTTACTCA 1080
Db 1021 TAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGATAGCTGGGACATTTACTCA 1080
Qy 1081 GTGCTCGTGAATGACACCTCAGAAGGAAATGAGTGTCTATAGAGAGAGAGAGTGT 1140
Db 1081 GTGCTCGTGAATGACACCTCAGAAGGAAATGAGTGTCTATAGAGAGAGAGAGTGT 1140
Qy 1141 ACTGCCAGGCTTTTGACAGATGTAATTTCTCATTTAAATTAAGTTTCAAGTTTGGTTA 1200
Db 1141 ACTGCCAGGCTTTTGACAGATGTAATTTCTCATTTAAATTAAGTTTCAAGTTTGGTTA 1200
Qy 1201 AGTGG 1205
Db 1201 AGTGG 1205

RESULT 2
BD085731
LOCUS BD085731 1205 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085731

BD085731.1 GI:22631341
SYNTHETIC CONSTRUCT
SYNTHETIC CONSTRUCT
OTHER SEQUENCES; ARTIFICIAL SEQUENCES.
1 (bases 1 to 1205)
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
Genes from the 20q13 amplicon and their uses
Patent: JP 2001524802-A 7 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/7
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
CC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:1b4
CC n is A, C, G, T, or U
CC n is A, C, G, T, or U
FH Key Location/Qualifiers
FT modified base (916)
FT modified base (937).
FEATURES
source
1..1205
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 99.7%; Score 1201; DB 6; Length 1205;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCGCGGTGAGTCCGCGCCCGCCAGTCAGCTGACCGCTGACCTCGGGGGCTTCTCCACTATC 60
Db 1 GCGCGGTGAGTCCGCGCCCGCCAGTCAGCTGACCGCTGACCTCGGGGGCTTCTCCACTATC 60
Qy 61 GCTTACCTACCTCCTCTGACGAAACCGCGGATATGGCTCGCGCTGTCGCCCGCGCGC 120
Db 61 GCTTACCTACCTCCTCTGACGAAACCGCGGATATGGCTCGCGCTGTCGCCCGCGCGC 120
Qy 121 ATTTCTCTCCCGCTGCTTCCCTTCTCTGCGCTTCTGCTCTCTCCGCTCCGATGCG 180
Db 121 ATTTCTCTCCCGCTGCTTCCCTTCTCTGCGCTTCTGCTCTCTCCGCTCCGATGCG 180
Qy 181 GGCAGCGCCTGCACACCAAGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCA 240
Db 181 GGCAGCGCCTGCACACCAAGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGTCA 240
Qy 241 TTCCCAAAGCAAGTTCGTCTGGTGAAGTTCGACACCCAGTACCCCTACGGTGAAGGA 300
Db 241 TTCCCAAAGCAAGTTCGTCTGGTGAAGTTCGACACCCAGTACCCCTACGGTGAAGGA 300
Qy 301 GGATGAGTTCAAGGCTCTTCTGAAACCTCGGCTCCAGCGATGATCTCTTGTCGTCAGAG 360
Db 301 GGATGAGTTCAAGGCTCTTCTGAAACCTCGGCTCCAGCGATGATCTCTTGTCGTCAGAG 360
Qy 361 GTGGGATCTCAGATTTATGTGACAACTGAACTGAGCTGAGTGAAGAAATACAACTGG 420
Db 361 GTGGGATCTCAGATTTATGTGACAACTGAACTGAGCTGAGTGAAGAAATACAACTGG 420
Qy 421 ACAAAGAGAGCTACCCATCTTCTAGCTCTTCCGGGATGGGACTTTGAGAACCCAGTCCC 480
Db 421 ACAAAGAGAGCTACCCATCTTCTAGCTCTTCCGGGATGGGACTTTGAGAACCCAGTCCC 480
Qy 481 ATACACTGGGCGAGTTAGTTGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGGTCTTAC 540
Db 481 ATACACTGGGCGAGTTAGTTGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGGTCTTAC 540

misc_feature	/note="AluSx repeat: matches 1. .302 of consensus" complement(2259. .2786) /note="match: GSS: Em:AQ549766" 2391. .2410	repeat_region	5982. .5993 /note="2.0 copies 6 mer CATTCT 24% conserved" 6019. .6511 /note="match: GSS: Em:AQ473868" 6490. .6500
repeat_region	/note="4.0 copies 5 mer CAAAA 40% conserved" 2417. .2432	repeat_region	/note="3.7 copies 3 mer TGC 22% conserved" complement(6634. .6722)
repeat_region	/note="5.3 copies 3 mer AAT 32% conserved" 2455. .2464	repeat_region	/note="MIR repeat: matches 11. .104 of consensus" 6740. .7042
repeat_region	/note="10.0 copies 1 mer T 20% conserved" 2482. .2491	repeat_region	/note="AluJb repeat: matches 1. .302 of consensus" 7023. .7042
repeat_region	/note="3.3 copies 3 mer TTA 20% conserved" 2542. .2561	repeat_region	/note="20.0 copies 1 mer A 40% conserved" 7164. .7177
repeat_region	/note="2.2 copies 9 mer TTGTGAGGC 31% conserved" 2544. .2657	repeat_region	/note="3.5 copies 4 mer TTAT 28% conserved" complement(7167. .7480)
misc_feature	/note="MER69 repeat: matches 17. .147 of consensus" complement(2592. .2794) /note="match: GSS: Em:AQ467488" 2685. .2731	repeat_region	/note="AluSp repeat: matches 1. .298 of consensus" 7240. .7273
repeat_region	/note="MER69 repeat: matches 2476. .2523 of consensus" 2803. .3228	repeat_region	/note="2.0 copies 17 mer TGGCAAACTTCCTCC 68% conserved" complement(7488. .7565)
misc_feature	/note="match: GSS: Em:AQ332266" 2859. .2943	repeat_region	/note="MIR repeat: matches 74. .156 of consensus" 7488. .7498
repeat_region	/note="MIR repeat: matches 61. .147 of consensus" 3001. .3014	repeat_region	/note="3.7 copies 3 mer ATT 22% conserved" 7607. .7616
repeat_region	/note="2.0 copies 7 mer AAAATGG 28% conserved" 3389. .3510	repeat_region	/note="2.5 copies 4 mer AAAC 20% conserved" complement(7623. .7917)
repeat_region	/note="MIR repeat: matches 2. .119 of consensus" 3398. .3409	repeat_region	/note="AluY repeat: matches 1. .295 of consensus" 7623. .7636
repeat_region	/note="2.0 copies 6 mer GTAATG 24% conserved" 3511. .3791	repeat_region	/note="14.0 copies 1 mer T 28% conserved" 8078. .8092
repeat_region	/note="AluY repeat: matches 1. .311 of consensus" 3621. .3647	repeat_region	/note="3.0 copies 5 mer GAAAA 30% conserved" complement(8471. .8780)
repeat_region	/note="1.9 copies 14 mer CCTGTAGTCCCAGG 54% conserved" 3762. .3791	repeat_region	/note="AluSg repeat: matches 1. .307 of consensus" 8482. .8494
repeat_region	/note="30.0 copies 1 mer A 60% conserved" 3792. .3917		
repeat_region	/note="MIR repeat: matches 119. .258 of consensus" complement(3955. .4264)	Qy	30 TGACCGGTGACTCGGGCGTTCTCCACTATCGCTTACTACTCTCCCTCTGCGAGAACCCG 89
repeat_region	/note="AluSx repeat: matches 1. .309 of consensus" 3955. .3981	Db	93229 TGACCGGTGACTTGGGGCGTTCTCCACTACGGCTTCCCACTCCCGCTCCCGCTTCCCTCTCC 93170
repeat_region	/note="27.0 copies 1 mer T 27% conserved" 3991. .4006	Qy	90 GCGATATGGCTGCGCTGTGCCCGCGCGCATTTCTTCTCCCGCTGCTT-CCCTTCTCC 148
repeat_region	/note="2.0 copies 8 mer TATTGTCTC 32% conserved" 4206. .4215	Db	93169 GCGATGTGGCTGCGCGCTGCGCGCTCTCTCTCCCTGCGCTTCCCTCTCC 93110
repeat_region	/note="2.5 copies 4 mer CTGC 20% conserved" 4275. .4323	Qy	149 TGGGCTTCTGCTCTCTCCGCTCGGATGCGGCGAGGCGCTGCACACCAAG-GCGCC 207
repeat_region	/note="9.8 copies 5 mer TTCT 54% conserved" 4279. .4324	Db	93109 TGGGCTTCTGCTCTCTCCGCTCCGATGCGAGCGGCTGCACACCAAGCGCGCC 93050
repeat_region	/note="46.0 copies 1 mer T 47% conserved" complement(4295. .4605)	Qy	208 TTCCCTGGATACGGTCACTTTCTACAAGGTCAITTCCTCCAAAAGCAAGTTCGTC-TGGTGA 266
repeat_region	/note="AluY repeat: matches 1. .311 of consensus" 4680. .4696	Db	93049 TTCCCTGGATACGGTCACTTTCTACAAGGTCAITTCCTCCAAAAGCAAGTTCGTC-TGGTGA 92990
repeat_region	/note="4.2 copies 4 mer ATGT 25% conserved" complement(4719. .5094)	Qy	267 AGTTTCGACACCCAGTACCCCTACCGTGAGAGAGAGAGATGAGTTCAAGCGCTTT-CTGAAA 325
repeat_region	/note="L1M4 repeat: matches 3083. .3493 of consensus" 4723. .4733	Db	92989 AGTTCAACACCCAGTACCGCTACGCTGAGAGACGAGATGAGTTCAAGCGCTTTTGTCTGAAA 92930
repeat_region	/note="2.2 copies 5 mer TAATT 22% conserved" 4879. .4890	Qy	326 ACTGGCTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT-GTGACA 384
repeat_region	/note="2.0 copies 6 mer TCTGCA 24% conserved" 5082. .5094	Db	92929 ACTGGCTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTATGTGTGACA 92870
repeat_region	/note="13.0 copies 1 mer T 26% conserved" complement(5159. .5450)	Qy	385 AGCTGAACATGGAGCTGAGTGAGAAATACAAAGCTGGACAAAGAGAGAGCTACCCA-TCCTTCT 443
repeat_region	/note="AluSg repeat: matches 10. .300 of consensus" 5159. .5178	Db	92869 AGCTGAACATGGAGCTGAGTGAGAAATACAAAGCTGCTCAAGAGAGCTACCCAGTCTTCT 92810
repeat_region	/note="20.0 copies 1 mer T 31% conserved" 5242. .5253	Qy	444 ACCTTCTCCGGATGGGACTTTGAGAACCCAGTCCCAATACACTGGGGCAGTT-AGGTTG 502
repeat_region	/note="2.0 copies 6 mer ACCTCC 24% conserved" 5505. .5514	Db	92809 ACCTTCTCCGGAGGGGACTTTGAGAACCCAGTCCCAATGATGACTGGGCGAGTTAAGGTTG 92750
repeat_region	/note="3.3 copies 3 mer TTA 20% conserved" 5514. .5548	Qy	503 GAGCATCAGCGCTGGCTGAAGGGGCAAGGGGTCTACTAGGTAT-----GCCTGG 554
repeat_region	/note="MIR repeat: matches 110. .144 of consensus" complement(5711. .5887)	Db	92749 GAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAGCTAGCTATGCCAGGTGCCCTGG 92690

QY 555 TGCCTGCTGTATACGACCCCTGCGGGAGTTTCATCAGCGCTCTTGTTGAGG-C 613
DB 92689 TGGCAGCTGTACATAGACACCTTGCC-AGGAGTTTCATCAGCGCTCTTGCGGTGAGGCC 92631
QY 614 CCCCAGGCGCTCTTGAAGCAGGCGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG 673
DB 92630 CGTAGGCGCTCTGAAGCAGGCGCAGGACCACTCTCAAGTGTGAAGGAGACTCAGAAG 92571
QY 674 AG-TGGCGCGAGCAATACCTGAAGATCATGTGGGAAGATCTTAGACAAGGGGAGCATTTC 732
DB 92570 AACTGGCGCGAGCAATACCTGAAGATCACAGGGAAGATCTTAGACTAGCAGGAGCTTC 92511
QY 733 CAGCATCAGATGACAGGATCGCCAGCTGATTGAGAGACAAAGATGATGACGCA 792
DB 92510 CAGCATGAGATGACATGGATCTCCAGCTGATTGAGAACAAAGTTGAGTGACG--GGAA 92453
QY 793 GAAGGAGGAGCTCCAGAAGAGCTTAAACATCTGCTGCTTCCAGAGAGAGGGGGCCGA 852
DB 92452 GAAGGAGGAGCTCCAGAAGAGCTTAAACATCTGCTTCCAGAGAGAGGGGGCTGA 92393
QY 853 GAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 912
DB 92392 GAAAGGAGGTTGTAAAAGGCTCTCTATGTTTCCAGGGTTTGGTAGGGTGGGGAGG 92333
QY 913 GGAAGTTAACTGCTGCTGTGATCCCTGTGTGAATATAGGGGGGMSKGGGAAAAGW 972
DB 92332 GGAGAGTTAACTGCTGCTGTGATCCCTGTGTGAATATAGGGGGTGGTGGGAAAAGT 92273
QY 973 GGTACTACCCAGCAATCTGAGCCTGAGTATGCTGACATGCTGACATGATGATCAATGACCAT 1032
DB 92272 GTTACTAACTTACAAATCTGAGCTCTGAGTATGCTGACATGCTGACATGATGATGACCAT 92213
QY 1033 GCTTGGGATGCTCTAGCTGCTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGA 1092
DB 92212 GCTTGGGATGCTCTAGCTGCTGCTGGGATAGCTGGAGCACTT-CTCAGGTGGCTGGTGA 92154
QY 1093 AATGACCTCAGAAGGAATGAGTCTATAGAGGAGAGAGAGTACTGCCCCAGTTC 1152
DB 92153 AATGACCTCAGAAGGAATGAGTCTATAGAGGAGAGTACTGCCCCAGTTC 92094
QY 1153 TTGACAGATGAATTTCTCATTAATTAAGTTTTCAGTGTGTTTGGTTAAGTGG 1205
DB 92093 TTGACAGTGAATTTGATTAATTAAGTTTTCAGTGTGTTTGGTTAAGTGG 92041

RESULT 4

AC073575/c
LOCUS Homo sapiens 12 PAC RPI-267L14 (Roswell Park Cancer Institute Human
DEFINITION PAC Library) complete sequence.

AC073575
AC073575.30 GI:27375026

HTG.
Homo sapiens (human)

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

AC073575
Homo sapiens

Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, K., Hart, M., Havlak, P., Haves, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lee, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louleeged, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marwede, I., Martin, R., Martindale, A., Martinez, E., Massey, E., Maxwell, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metzger, M., Miller, A., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Neal, D., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokoko, S., Ogih, M., Okuwon, G., Oragunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojebokan, I., Rolfe, M., Ruiz, S., Savary, G., Scher, S., Scott, G., Shen, H., Shim, C., Shooshtari, N., Sisson, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umami, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y., Wu, Y.P., Zhou, J., Zorrilla, S., Zorrilla, S., Kucherlapati, R., Weinstock, G. and Gibbs, R.

Unpublished
2 (bases 1 to 173718)
Worley, K.C.
Direct Submission
Submitted (25-JUN-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 173718)
Worley, K.C.
Direct Submission
Submitted (27-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 173718)
Worley, K.C.
Direct Submission
Submitted (27-DEC-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 27, 2002 this sequence version replaced gi:2213011.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences. Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum


```
QY 972 WGGTACTAACCCACGATCTGAGCCCTGAGTATGCTGAGACATTGATGCTAACATGACCA 1031
Db 158327 TGGTACTAACCCACGATCTGAGCCCTGAGTATGCTGAGACATTGATGCTAACATGACCA 158268

QY 1032 TGGTGGGATGTCTCTAGTGGTGTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTG 1091
Db 158267 TGGTGGGATGTCTCTAGTGGTGTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTG 158208

QY 1092 AAATGACACCTCAGAAGAAATAGTGTCTATAGAGAGAGAGAGAGTGTACTGCCCCAGGT 1151
Db 158207 AAATGACACCTCAGAAGAAATAGTGTCTATAGAGAGAGAGAGAGTGTACTGCCCCAGGT 158148

QY 1152 CTTTGACAGATGTAATCTCATTTCAATTAAGTTTCAGTGTGTTGTTAAAGTGG 1205
Db 158147 CTTTGACAGATGTAATCTCATTTCAATTAAGTTTCAGTGTGTTGTTAAAGTGG 158094

RESULT 5
HSERP28 HSERP28 892 bp mRNA linear PRI 07-AUG-1998
LOCUS Homo sapiens mRNA for Erp28 protein.
DEFINITION X94910
ACCESSION X94910.1 GI:3413292
VERSION ER protein; Erp28 protein; protein disulfide isomerase.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ferrari, D.M., Nguyen Van, P., Kratzin, H.D. and Soling, H.D.
TITLE Erp28, a human endoplasmic-reticulum-luminal protein, is a member
of the protein disulfide isomerase family but lacks a CXXC
thioredoxin-box motif
JOURNAL Eur. J. Biochem. 255 (3), 570-579 (1998)
MEDLINE 98409275
PUBMED 9738895
REFERENCE 2 (bases 1 to 892)
AUTHORS Ferrari, D.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-1996) D.M. Ferrari, Zentrum Innere Medizin, Abt.
Klinische Biochemie, Robert-Koch-Str. 40, D-37075 Goettingen, FRG
COMMENT On Aug 12, 1998 this sequence version replaced gi:1154677.
Related to protein disulfide isomerase (PDI).
FEATURES
source
1..892
/organism="Homo sapiens"
/mol_type="mRNA"
/isolate="clone 28-3"
/species="human"
/db_xref="taxon:9606"
/clone="E.coli XL1-Blue"
/cell_type="hepatocytes"
/tissue_type="liver"
/clone_lib="lambda gt-11"
/dev_stage="adult"
/lab_host="E.coli"
12..797
/function="ER protein"
/notes="Similar to ULA_5 product, AC P30040"
/codon_start=1
/evidence=experimental
/product="Erp28"
/protein_id="CAA64397.1"
/db_xref="GI:3413293"
/db_xref="GOA:P30040"
/db_xref="UniProt/Swiss-Prot:P30040"
/translation="MAAAVPRAAFLSPLLGLFLLLSAPHGSGSLHTKALPDLTV
TFYKVIKSKFVLVKFDYQYFYGKQDEFKRLAENSASSDLLVAEYISDYGDKLNN
ELSKYKLDKESYFVFLFRDGFENFVPTGAVKVGAIQRLKQGVYLGMPGLPV
YDALAGEFIRASGVEARQALLKQODNLSSVKETQKKWAEOLYKIMKILQDQDFPA
SEMTRIALIEKNMSDGGKKEELQKSLNLTAPQKGAKEEL"
12..106
sig_peptide
```

LOCUS CQ728979 889 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 14913 from Patent WO02068579.
ACCESSION CQ728979
VERSION CQ728979.1 GI:42299268
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 14913 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..889
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 61.3%; Score 738.6; DB 6; Length 889;
Best Local Similarity 97.9%; Pred. No. 2.9e-198;
Matches 870; Conservative 3; Mismatches 4; Indels 12; Gaps 12;
Qy 86 CCCGCGATATGGCTCGCGCTGCGCCCGCGCGGATTTCTCCCGCTGCTT-CCCTT 144
Db 1 CCCGCGATATGGCTGCGCGCTGCGCCCGCGCGGATTTCTCCCGCTGCTTCCCTT 60
Qy 145 CTCCTGGGCTTCCTGCTCTCTCCGCTCCGATGCGCGGCGCTGCACACCAA-GGC 203
Db 61 CTCCTGGGCTTCCTGCTCTCTCCGCTCCGATGCGCGGCGCTGCACACCAAGGC 120
Qy 204 GCCCTTCCCTGGATACCGGTCACTTTTACAAGGTCAATCCCAAAGCAAGTTGCTC-TG 262
Db 121 GCCCTTCCCTGGATACCGGTCACTTTTACAAGGTCAATCCCAAAGCAAGTTGCTC-TG 180
Qy 263 GTGAAGTTGACACCCAGTACCCCTACCGTTCAGTGAAGCAGGATGATTTCAAGCGCTTT-CT 321
Db 181 GTGAAGTTGACACCCAGTACCCCTACCGTTCAGTGAAGCAGGATGATTTCAAGCGCTTTCTGCT 240
Qy 322 GAAACTCGGCTTCAGCGATGATCTCTGTGGCAGAGGTGGGATCTCAGATTAT-GT 380
Db 241 GAAACTCGGCTTCAGCGATGATCTCTGTGGCAGAGGTGGGATCTCAGATTATGCT 300
Qy 381 GACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGAACAAAGAGAGCTACCCA-TC 439
Db 301 GACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGAACAAAGAGAGCTACCCAGTC 360
Qy 440 TTCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTT-AG 498
Db 361 TTCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTTAAG 420
Qy 499 GTTGAGGCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTACCTAGGTATGCTGG-TGC 557
Db 421 GTTGAGGCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTACCTAGGTATGCTGGTTGC 480
Qy 558 CTGCTGTATACGACGCCCTGCGCGGGAGTTTCATCAGGGCTCTGTGTGTGAGG-CCGC 616
Db 481 CTGCTGTATACGACGCCCTGCGCGGGAGTTTCATCAGGGCTCTGTGTGTGAGGCGCCGC 540
Qy 617 CAGGCGCTTGAAGCAGGGGCAAGTACCTCTCAAGTGTGAAGGAGACTCAGAG-AG 675
Db 541 CAGGCGCTTGAAGCAGGGGCAAGTACCTCTCAAGTGTGAAGGAGACTCAGAGGAAG 600
Qy 676 TGGGCGGAGCAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTT-CCA 734
Db 601 TGGGCGGAGCAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTTCCA 660
Qy 735 GCATCAGAGATGACACGATCCCGAGGCTGATTGAGAGAAACAGATGAGTGAC-GGAG 793
|||||

Db 661 GCATCAGAGATGACACGATCGCCAGGCTGATGAGAGAAACAAGATGAGTGACGGGAAG 720
Qy 794 AAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCTCTCCAGAAGAAAGGGGGCCGAG 853
Db 721 AAGGAGGAGCTCCAGAAGAGCTTAAACATCTGACTGCTCTCCAGAAGAAAGGGGGCCGAG 780
Qy 854 AAGGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGG 913
Db 781 AAGGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGG 840
Qy 914 GANAGTTAACTGCTGCTGTGANTCCCTTGTGGAATATAAGGGGGYMS 962
Db 841 GAGAGTTAACTGCTGCTGTGANTCCCTTGTGGAATATAAGGGGGTAG 889
RESULT 7
LOCUS CR541667 786 bp mRNA linear PRI 29-JUN-2004
DEFINITION Homo sapiens full open reading frame cDNA clone RZPD0834E0727D for
gene C12orf8, chromosome 12 open reading frame 8; complete cds,
incl. stopcodon.
ACCESSION CR541667
VERSION CR541667.1 GI:49456294
KEYWORDS Full ORF shuttle clone, Gateway(TM), complete cds.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 786)
Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar.W.,
Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
TITLE Cloning of human full open reading frames in Gateway(TM) system
entry vector (pDONR201)
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 786)
Halleck,A., Ebert,L., Mkoundinya,M., Schick,M., Eisenstein,S.,
Neubert,P., Kstrang,K., Schatten,R., Shen,B., Henze,S., Mar.W.,
Korn,B., Zuo,D., Hu,Y. and LaBaer,J.
TITLE Direct Submision
JOURNAL Submitted (28-JUN-2004) RZPD Deutsches Ressourcenzentrum fuer
Genomforschung GmbH, Im Neuenheimer Feld 580, D-69120 Heidelberg,
Germany
COMMENT
RZPD; RZPD0834E0727D, ORFNO 3284
This clone is available from RZPD;
www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPD0834E0727D RZPDLIB;
Human Full ORF Clones Gateway(TM) - RZPD (kan-resist.) RZPD LIB No.
834
www.rzpd.de/cgi-bin/products/showLib.pl.cgi/response?libNo=834
www.rzpd.de/products/orfclones/
Contact: Inge Arlart
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 100
Fax: +49 30 32639 111
www.rzpd.de
This clone is available from RZPD;
Contact RZPD (customer.service@rzpd.de) for further information.
This CDS clone is part of a collection of human full ORF clones
jointly established and verified by the Harvard Institute of
Proteomics (HIP) and RZPD.
The CDS has been inserted into pDONR201 via a BP Clonase(TM)
reaction. Additional sequence has been added in front of the start
codon: att. AAAAAA GCA GGC TCC ACC (ATG).
The stopcodon is followed by the 3' att site: GACCCAGCTTCTT..att
The clone is validated by full sequence check.
Compared to the reference sequence NM_006817 (GI:13124889) we did
not find any amino acid exchanges.
Clone distribution: http://www.rzpd.de/products/orfclones/
FEATURES
source Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"

```
/db_xref="taxon:9606"
/clone="RZPD0834E072D"
/clone_lib="Human Full ORF Clones Gateway(TM) - RZPD"
/lab_host="DH5Alpha"
/note="Vector: pDONR201, Site_1: attP1; Site_2: attP2"
1..786
/gene="C12orf8"
CDS
1..786
/gene="C12orf8"
/codon_start=1
/protein_id="CAG46468.1"
/db_xref="GI:49456295"
translation="MAAAVPRAAFSPLLPLLLGLLSAPHGGSLHTKCALPLDVT
TFYKVPKSVLKFDFQYPYGEKQDFKLAENSASSDLLVAEIVGISYDGKLNW
ELSEYKLDKESYPFVFLFRDGFENFVFTGAKVGAIQRLWAGQGVILMPGCLPV
YDALAGEFIRASGVEARQALLKQGDNLSSVKETQKKWAEQYLMKIMGILKQGGDFPA
SEMTRIARLIEKNRMSDGKBELOKSLNLTAFQKGAKEEL"
ORIGIN
Query Match 52.7%; Score 635.6; DB 9; Length 786;
Best Local Similarity 98.0%; Pred. No. 5.2e-169;
Matches 770; Conservative 0; Mismatches 4; Indels 12; Gaps 12;
Qy 95 ATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCTTCTCTCTGGGC 153
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 1 ATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCTGCTTCTCTCTGGGC 60
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 154 TTCTGTCTCTCTCGCTCCGATCGGCGGAGCGGCTGACACAA-GGGCGCCCTTCCC 212
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 61 TTCTGTCTCTCTCGCTCCGATCGGCGGAGCGGCTGACACAAAGGCGCGCTTCCC 120
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 213 CTGATACGCTCACTTTCTACAGGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTC 271
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 121 CTGATACGCTCACTTTCTACAGGTCAATTCCTCCAAAGCAAGTTCGTC-TGGTGAAGTTC 180
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 272 GACACCCAGTACCCCTAGCTGAGAGCAGATGATTCAGAGCTTT-CTGAAAACCTCG 330
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 181 GACACCCAGTACCCCTAGCTGAGAGCAGATGATTCAGAGCTTTCTGCTGAAAACCTCG 240
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 331 GCTTCCAGCATGATCTCTTGGTGACAGAGTGGGATCTCAGATTAT-GTGCAAGCTG 389
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 241 GCTTCCAGCATGATCTCTTGGTGACAGAGTGGGATCTCAGATTATGTGTGCAAGCTG 300
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 390 AACATGAGCTGAGTGAGAAATACAGCTGACAAAGAGCTACCCA-TCTTCTACCTC 448
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 301 AACATGAGCTGAGTGAGAAATACAGCTGACAAAGAGCTACCCAAGTCTTCTACCTC 360
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 449 TTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTT-AGTTTGGAGCC 507
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 361 TTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTTAAGTTGGAGCC 420
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 508 ATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTATGCTCTGG-TGCTGCTGTA 566
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 421 ATCCAGCGCTGGTGAAGGGCAAGGGGTCTACCTAGGTATGCTCTGGTTGCTGCTGTA 480
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 567 TAGACGCGCTGGCGGGGAGTTATCAGGCGCTCTGTTGTGAGG-CCGCAAGCCCTC 625
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 481 TACGATGCTTGGCGGGGAGTTATCAGGCGCTCTGTTGTGAGGCGCCGCGAGCCCTC 540
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 626 TTGAAGCAGGGCAGAGTAACCTCTCAGTGTGAGGAGCTCAGAG-AGTGGCCGAG 684
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 541 TTGAAGCAGGGCAGAGTAACCTCTCAGTGTGAGGAGCTCAGAGAGAGTGGGGCAG 600
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 685 CAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAGAG 743
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 601 CAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAGAG 660
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 744 ATGACACGGATCGCCAGCTGATTGAGAAGAACAGATGAGTAC-GGCAGAGGAGAG 802
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 661 ATGACACGGATCGCCAGCTGATTGAGAAGAACAGATGAGTACGCGGAAGAGGAGAG 720
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 803 CTCAGAGAGCTTAACATCTGCTCTTCAGAGAGGGGGCGAGAAAGAGAG 862
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

FEATURES

```
Db 721 CTCGAGAGAGCTTAACATCCTGACTGCTTCAGAGAAAGGGGCCGAGAAAGAGAG 780
Qy 863 CTGTAA 868
Db 781 CTGTAA 786

RESULT 8
BC017125 1258 bp mRNA linear ROD 30-JUN-2004
LOCUS Mus musculus endoplasmic reticulum protein 29, mRNA (cdna clone
DEFINITION MGC:27552 IMAGE:4481563), complete cds.
ACCESSION BC017125
VERSION BC017125.1 GI:16877775
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1258)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heish,F.,
Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Scapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A.; McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,D.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnierch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
GENERATION AND INITIAL ANALYSIS OF MORE THAN 15,000 FULL-LENGTH
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1258)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skaleka,
Duane Smalls, Jeff Stott, Miranda Tsai, George Yang, Jackie
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 38 Row: a Column: 13
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19526462.
Location/Qualifiers
```


QY 56 CTATCGCTTACCTACCTCCCTCTGAGGAACCCGGATATGGCTGCGCTGTGCCCCGC 115
Db 5 CTGTAGCTTGCCCATCTCCGCTCCACATCCGGGCTGATGGCTGCCCGCTTCTGCG 64
QY 116 GCCGATTTCTCTCCCGCTGCT-TCCCTTCTCTCCGGCTTCTGCTCTCTCCGCTCG 174
Db 65 GCGGTTTGGCTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124
QY 175 CATGGCGGAGGGCTGACACCAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTAC 233
Db 125 CATGGCGGAGGGCTGACACCAAAGGGCGGCTTCCCTTGGACACAGTCACTTTCTAC 184
QY 234 AAGGTCATTCCCAAAAGCAAGTTCTGCT-TGGTGAAGTTCCGACACCAAGTACCCCTACGGT 292
Db 185 AAGGTCATTCCCAAAAGCAAGTTCTGCTTGGTGAAGTTCCGACACCAAGTACCCCTATGGA 244
QY 293 GAGAAGCAGGATGAGTTCAAGCGTCT-TCTGAAAACTCGGCTTCCAGCGATGATCTCTTG 351
Db 245 GAGAAGCAGGATGAGTTTAAAGCGTCTGGCTGAGAACTCAGCCTCCAGCGATGATCTTTG 304
QY 352 GTGGCAGAGTGGGATCTCAGATTAT-GTGAACAAGCTGAACATCGAGCTGAGTGAGAAA 410
Db 305 GTGGCAGAGTGGGATCTCAGATTATGCTGCAAGCTGAACATCGAGCTGAGTGAGAA 364
QY 411 TACAAGCTGGACAAAGAGAGCTACCCA-TCTTCTACCTCTCCGGATGGGACTTTGAG 469
Db 365 TACAAGCTGGACAAAGAGAGCTTACCAGTCTTCTACCTCTCCGGATGGGACTTTGAG 424
QY 470 AACCCAGTCCCATACACTGGGCGAGT-AGGTTGGAGGCATCCAGCGCTGGCTGAAGGG 528
Db 425 AATCTGTCTCCCATACAGGGGCAAGTAAAGTTGGAGCCATCCAGCGTGGCTCAAGGG 484
QY 529 CAAGGGGTCTACTAGGTATGCTG-TCGCTGCTGTATACGACGCCCTGGCGGGAG 587
Db 485 CAGGGAGTCTATCTGGCATGCTGGATGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
QY 588 TTCAATCAGGCTCTGCTGTGAGG-CGCGCAGGCGCTTGAACAGGCGGCAGATAC 646
Db 545 TTCAATCAGGCGCTCCAGCAGAGAGCGCCGCGCAGGCATCTGTAACAGAGGGCAGATGC 604
QY 647 CTCTCAAGTGTGAAGGAGACTCAGAA-AGTGGGCGGAGCAATACCTGAAGATCATGGG 705
Db 605 CTCTCAGGTGTGAAGGAGACACAAAGAGTGGGCGAGTCACTGAGATCATGGG 664
QY 706 AAGATCTTAGACCAAGGGAGCACTT-CCAGCATCAGAGATGACACGATGCCAGGCTG 764
Db 665 AAGATCTTGGACCAAGTGAAGACTTCCGGCTCCGAGCTGGCGCGATCAGTAAGCTC 724
QY 765 ATTGAGAAGNACAAGATGATGATGACCGCAGAGGAGGCTCCAGAGAGCTTAAACATCC 824
Db 725 ATTGAGAACNAGATGATGAGG--GTAAGAAGGAAGAGCTGCAGAGGAGCCTCAACATCC 782
QY 825 TGACTGCTCTCCAGAAAGGGGCGGAGAAAGAGAGAGCTGTAAAGAGCTGTCTGTGAT 884
Db 783 TCACCGCTTCCGAAAGAGGCGCGGAGAGAGAGAGCTGT-----GAGGGACACAGG 836
QY 885 TTTCAGGGTTTGTGGGGGTAGGAGGGGAGNAGTTAACTGCTGCTGTGTGANTCCCTTG 944
Db 837 CTGCTAGGTTTGTGAGGGGGGAGG--AGAGCGGCTGCTGACTGTGGGGCCCTGT 894
QY 945 TGGAAATTAAGGGGGMKGGGAAAGGAGTAAACCAAGTCTGAGCGCTGAGTAT 1004
Db 895 GGG--TGGAGGGGGCAGTGGAGCAAGCAGTCTTGGAGCAGAGAGCTGTGGCCCCAG--T 950
QY 1005 GCCTGACATTCATGCTAACATGACCATGCTTGGGATGCTCTAG 1049
Db 951 GCCTGACACTGAGGCTGCTGAGACCATACCTCGACATCTCTTG 995

RESULT 10

RNU36482

LOCUS

DEFINITION Rattus norvegicus endoplasmic reticulum protein Erp29 precursor,

RNU36482

1139 bp

mRNA

linear

ROD 29-APR-1998

ACCESSION

U36482

VERSION

U36482.1

KEYWORDS

GI:2317799

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

REFERENCE

1 (bases 1 to 1139)

Mkrtchian,S., Fang,C., Hellman,U. and Ingelman-Sundberg,M.

A stress-inducible rat liver endoplasmic reticulum protein, Erp29

Eur. J. Biochem. 251 (1-2), 304-313 (1998)

JOURNAL

98151262

MEDLINE

9492298

REFERENCE

2 (bases 1 to 1139)

Mkrtchian,S., Fang,C., Hellman,U. and Ingelman-Sundberg,M.

Direct Submission

TITLE

Submitted (18-SEP-1995) Souren Mkrtchian, Dept. of Medical

JOURNAL

Biochemistry and Biophysics, Karolinska Institute, Doktorsträngen

12B, Stockholm S-171 77, Sweden

COMMENT

On Aug 11, 1997 this sequence version replaced gi:1172121.

FEATURES

Location/Qualifiers

1..1139

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/tissue type="liver endoplasmic reticulum"

44..826

/note="translation of this region is similar to unknown

protein from 2D-PAGE of liver tissue, Swiss-Prot Accession

Number P30040"

/codon_start=1

/product="endoplasmic reticulum protein Erp29 precursor"

/protein_id="AAC15239.1"

/db_xref="GI:2317800"

/translation="MAAIVPGAVSLPILLSVLGLLLSLAPHSGLHTKGLPLDVT

TFYKIVPSKFLVKFYDTPYQGFDEFKRLAENSASSDLLLVAEIGISIDGKLNM

ELSKYKLDKESYFVFLFRDGFENFVPGVAVKGAIRWLKQGVLYGMPCGLPA

YDALAGQFIASREARQAILKQGDGLSGVKETDKWASQYLKIMKILDKQGEDFPA

SELARIKLIENKMSKGBELQSLNLTAFRKKGAKEEL"

44..139

/evidence=experimental

140..823

/product="endoplasmic reticulum protein Erp29"

812..823

/note="putative endoplasmic reticulum retrieval signal

(KEEL)"

polyA_signal

1096..1110

ORIGIN

Query Match

41.2%; Score 496.6; DB 10; Length 1139;

Best Local Similarity

78.8%; Pred. No. 1.7e-129;

Matches 792; Conservative

1; Mismatches 187; Indels 25; Gaps 16;

QY 56 CTATCGCTTACCTACCTCCCTCTGAGGAACCCGGATATGGCTGCGCTGTGCCCCGC 115

Db 5 CTGTAGCTTGCCCATCTCCGCTCCACATCCGGGCTGATGGCTGCCCGCTTCTGCG 64

QY 116 GCCGATTTCTCTCCCGCTGCT-TCCCTTCTCTGGGCTTCTGCTCTCTCCGCTCG 174

Db 65 GCGGTTTGGCTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 124

QY 175 CATGGCGGAGGGCTGACACCAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTAC 233

Db 125 CATGGCGGAGGGCTGACACCAAAGGGCGGCTTCCCTTGGACACAGTCACTTTCTAC 184

QY 234 AAGGTCATTCCCAAAAGCAAGTTCTGCT-TGGTGAAGTTCCGACACCAAGTACCCCTACGGT 292

Db 185 AAGGTCATTCCCAAAAGCAAGTTCTGCTTGGTGAAGTTCCGACACCAAGTACCCCTATGGA 244

QY 293 GAGAAGCAGGATGAGTTCAAGCGTCT-TCTGAAAACTCGGCTTCCAGGATGATCTCTTG 351


```
Db 245 GAGAAGCAAGATGAGTTTAAAGCGTCTGGCTGAGAACTCAGCCTCAGCGATGATCTCTTG 304
Qy 352 GTGGCAGAGGTGGGATCTCAGATTAT-GTGACAAAGCTGAACATGGAGCTCAGTGAGAAA 410
Db 305 GTGGCAGAGGTGGGATCTCAGACTATGCTGACAAAGCTGAACATGGAGCTCAGTGAGAA 364
Qy 411 TACAAGCTGGAACAAGAGAGCTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTTGA 469
Db 365 TACAAGCTGGAACAAGAGAGCTACCCAGTCTTCTACCTCTTCCGGGATGGGACTTTGAG 424
Qy 470 AACCAGTCCCATATACACTGGGGCAGTT-AGTTGGAGCCATCCAGCGTGGCTGAAGGG 528
Db 425 AATCTGTCCCATACAGCGGGGCAAGTTAAAGTTGGAGCCATCCAGCGCTGGCTCAAGGG 484
Qy 529 CAAGGGGTCTACCTAGGTATGCTGG-TGCTGCTCTGTATACGACGCCCTCGCCGGGAG 587
Db 485 CAGGAGTCTATCTGGGCATGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
Qy 588 TTCTATCAGGGCTCTGGTGTGGAGG-CGCGCAGAGCCCTCTTTGAAGCAGGGGCAAGATAAC 646
Db 545 TTCTATCAGGGCTCTCAGCAGAGAGGCGCGCCAGGCGCATCTCTGAAACAGGGGCGAGTGGC 604
Qy 647 CTCTCAAGTGTGAAGAGACTCAGAG-AGTGGGCGGAGCAATACCTGAAGATCATGGG 705
Db 605 CTCTCAGGTGTGAAGGAGACAGACAAGAAGTGGGCCAGTCAAGTACCTGAAGATCATGGG 664
Qy 706 AAGATCTTAGACCAAGGGGAGCATT-CCAGCATCAGAGATGACACGGATCGCCAGGCTG 764
Db 665 AAGATCTTAGACCAAGGGTGAAGACTTCCCGGCTCCGAGCTGGCCGGATCAGTAAGCTC 724
Qy 765 ATTGAGAAACAAGATGAGTCAAGGCAAGAGGAGCTCCAGAGAGCTTAAACATCC 824
Db 725 ATTGAGAAACAAGATGAGTGAGG-GTAAAGAGGAAGAGCTCAGAGGAGCTCAACATCC 782
Qy 825 TGACTGCTTCCAGAAAGGGGCGGAGAAAGAGGAGCTGTAAAGAGCTGTCTGTAT 884
Db 783 TCACGCGCTTCGCAAGAAGGCGCGGAGAGGAGCTGT-----GAGGGCACCCAGG 836
Qy 885 TTTCCAGAGTTTGGTGGGGTGGGAGGAGGAGTAACTGCTGGCTGTGANTCCCTTG 944
Db 837 CTGCTAGGTTTGTGAGGGCGGGAGGG--AGAGCGCTGCTGCTGCTGCTGGGCGCTGT 894
Qy 945 TGGATATAAGGGGYSKGGGAAAGWGTTACTAAACCCACGATTTCTGAGCCCTGAGTAT 1004
Db 895 GGG--TGGAGGGGCGAGTGGAGCAAGAGCAGTCTTGAGCCAGAGACCTGTGCCCGAG--T 950
Qy 1005 GCCTGGACATGATGCTTAACATGACCATGCTTGGGATGCTCTAG 1049
Db 951 GCCTGGACACTGAGGCTGCTGAGACCATACCTGGACATCTCTTG 995

RESULT 11
RNERP29PR RNERP29PR 1084 bp mRNA linear ROD 10-MAR-1997
LOCUS R.norvegicus mRNA for Erp29 protein.
DEFINITION Y10264
ACCESSION Y10264
VERSION Y10264.1 GI:1772593
KEYWORDS ER protein; Erp29. (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1
REFERENCE Demmer,J., Zhou,C. and Hubbard,M.J.
AUTHORS Molecular cloning of Erp29, a novel and widely expressed resident
TITLE of the endoplasmic reticulum
JOURNAL FEBS Lett. 402 (2-3), 145-150 (1997)
MEDLINE 97188599
PUBMED 9037184
REFERENCE 2 (bases 1 to 1084)
AUTHORS Hubbard,M.J.
TITLE Direct Submission
```

```
JOURNAL Submitted (24-DEC-1996) M.J. Hubbard, University of Otago,
COMMENT Department of Biochemistry, Box 56 Dunedin, 9001, NEW ZEALAND
FEATURES Related sequences: U36482 and X94910.
source Location/Qualifiers
1..1084
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="dental enamel epithelium (molar)"
/clone_lib="lambda ZAP Express"
/lab_stage="neonate"
/lab_host="E.coli"
15..797
/function="ER protein"
/codon_start=1
/evidence=experimental
/product="Erp29 precursor"
/protein_id="CAA71313.1"
/db_xref="GI:1772594"
/db_xref="GOA:P52555"
/db_xref="UniProt/Swiss-Prot:P52555"
/translation="MAAAVPGAVSLPILLSVLGLLLLSAPHGASGLHTKCALPLDTV
TFKVPKSFVLVFDFTQPYGKQDEFKRLAENSASSDILLVAEVGISDYGDKLMW
ELSEKYLDKESYFVFLFRDGFENFYPVPGAVKVGAIQRWLKQGVYILGPGCLPA
YDALAQFI EASREARQAILKQDGLSGVKETDKWASOYLKIMKILDOQGEDFPA
SELARISKLIENKMSGKBEQLRSLNILTAFRKKGAKEEL"
15..110
sig_peptide
/evidence=experimental
111..794
mat_peptide
/product="Erp29"
/evidence=experimental
783..794
misc_signal
/note="putative ER-retention signal (KEEL)"
1084..1089
polyA_signal
/evidence=not_experimental
ORIGIN
Query Match 39.8%; Score 479.4; DB 10; Length 1084;
Best Local Similarity 78.9%; Pred. No. 1.3e-124;
Matches 771; Conservative 1; Mismatches 179; Indels 26; Gaps 16;
Qy 84 AACCCGCGATATGGCTGCGCGCTGTGCCCCGCGCATTTCTCTCCCGCTGT-TCCCC 142
Db 4 AATCCGCGGTGATGGCTGCGCGCTTCTCTGCGCGCTTCTCTCCCGCTGTCTCG 63
Qy 143 TTCTCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 201
Db 64 TTCTCTGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 123
Qy 202 GCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCAATTTCCCAAAAGCAAGTTCTGTC- 260
Db 124 GCGCCCTTCCCTTGGACACAGTCACTTTCTACAGGTCAATTTCCCAAAAGCAAGTTCTGTC 183
Qy 261 TGGTGAAGTTTGACACCCAGTACCCCTACCGTGAAGAGCAGGATGAGTTCAAGCGTCT-T 319
Db 184 TGGTGAAGTTTGACACCCAGTACCCCTATGAGAGAAAGCAAGATGAGTTTAAAGCGTCTGG 243
Qy 320 CTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT- 378
Db 244 CTGAGAACTCAGCTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTATG 303
Qy 379 GTGACAACTGAACATGGAGCTGAGTGAGAAATACAGCTGACAAAGAGAGCTACCCA- 437
Db 304 GTGACAACTGAACATGGAGCTGAGTGAGAAATACAGCTGACAAAGAGAGCTACCCA 363
Qy 438 TCTTCTACCTTTCGCGGATGGGACTTTTGAGAACCCAGTCCCCATACACTGGGCGAGTT- 496
Db 364 TCTTCTACCTTTCGCGGATGGGACTTTTGAGAAATCTCTGTCCCATACAGCGGGGAGTTA 423
Qy 497 AGTTGGAGCCATCCAGGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTATGCTGTG-T 555
Db 424 AAGTTGGAGCCATCCAGGCTGGCTCAAGGGGCGAGGAGTCTATCTGGGCGATGCTGTGAT 483
```

```
QY 556 GCTGCTCTGTATACGACCCCTGGCCGGGAGTTTCATCAGGCGCTCTGTGTGGAG-CC 614
Db 484 GTCTGCTCTGTATACGACCCCTGGCCGGGAGTTTCATCAGGCGCTCTGTGTGGAG-CC 543
QY 615 GCCAGGCGCTCTGTGAAGCAGGCGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG- 673
Db 544 GCCAGGCGCTCTGTGAAGCAGGCGGCAAGATAAAGCTCTCAAGTGTGAAGGAGACTCAGAAG- 603
QY 674 AGTGGGCGGAGCATACCTGAAGATCATGGGGAAGATCTTGAACCAAGGGGAGCATT-C 732
Db 604 AGTGGGCGGAGCATACCTGAAGATCATGGGGAAGATCTTGAACCAAGGGGAGCATTCC 663
QY 733 CAGCATCAGAGATGACACCGATCGCCAGGCTGATGAGAGAACAGATGAGTGCAGCA 792
Db 664 CGGCTCTCGAGCTGGCCCGGATCAGTAAGCTCATGAGAACAGATGAGTGCAGCA 721
QY 793 GAAGGAGGAGCTCCAGAGAGCTTAAACATCTCTGACTGCTTCCAGAGAAAGGGGCGCA 852
Db 722 GAAGGAGGAGCTCCAGAGAGCTTAAACATCTCTGACTGCTTCCAGAGAAAGGGGCGCA 781
QY 853 GAAAGGAGGAGCTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 912
Db 782 GAAAGGAGGAGCTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 834
QY 913 GGAAGTAACTGCTGCTGGCTGTGANTCCCTTGTGGAATATAGGGGYSKGGGAAAGW 972
Db 835 G--AGAGCGCTGTGCTGACTGTGGGCGCTGTGGG--TGGAGGGGCGAGTGGAGCAAGC 890
QY 973 GGTACTTAACCCAGCTTCTGAGCCCTGAGTATGCTCTGAGTATGCTTAACATCACCATT 1032
Db 891 AGTCTGAGCAGAGACCTGTGCCCCGAG--TGCTTGGACACTGAGGCTGCTGAGACCAT 948
QY 1033 GCTTGGGATGCTCTAG 1049
Db 949 ACCCTGGACATCCTTGG 965

RESULT 12
CQ692426 438 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 37352 from Patent WO02070737.
ACCESSION CQ692426
VERSION CQ692426.1 GI:42233241
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Liaw,C.C., Marshall,W.E. and Zhang,H.
TITLE Compositions and methods relating to osteoarthritis
JOURNAL Patent: WO 02070737-A 37352 12-SEP-2002;
Chondrogene Inc. (CA)
FEATURES
source
1..438
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 33.6%; Score 404.8; DB 6; Length 438;
Best Local Similarity 97.0%; Pred. No. 1.9e-103;
Matches 425; Conservative 5; Mismatches 6; Indels 2; Gaps 2;
QY 691 CTGAAGATCATGGGAGAGCTTTAGACCAAGGGGAGCATT-CCAGCATCAGAGATGACA 749
Db 1 CTGAAGATCATGGGAGAGCTTTNACCAAGGGGAGGAGCTTCCAGCATCAGAGATGACA 60
QY 750 CGGATCGCCAGGCTGATTCAGAGAACCAAGATGAGTGCAG-GGCAGAGGAGGAGCTCCAG 808
Db 61 CGGATCGCCAGGCTGATTCAGAGAACCAAGATGAGTGCAGGAGGAGGAGGAGGAGCTCCAG 120
```

```
QY 809 AAGAGCTTAAACATCCTGACTGCTTCCAGAGAGGGGGCCGAGAAAGAGAGAGCTGTAA 868
Db 121 AAGAGCTTAAACATCCTGACTGCTTCCAGAGAGGGGGCCGAGAAAGAGAGAGCTGTAA 180
QY 869 AAAGGCTCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGANAGTTAACTGCT 928
Db 181 AAAGGCTCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGAGAGTTAACTGCT 240
QY 929 GGTGTGTANTCCCTTGTGGAATATAGGGGYSKGGGAAAGWGGTACTAAACCCACGAT 988
Db 241 GGTGTGTAGTCCCTTGTGGAATATAGGGGYSKGGGAAAGWGGTACTAAACCCACGAT 300
QY 989 TCTGAGCCCTGAGTATGCTGACATTTGATCTTAACATGACCATCTTGGGATGTCTTA 1048
Db 301 TCTGAGCCCTGAGTATGCTGACATTTGATCTTAACATGACCATCTTGGGATGTCTTA 360
QY 1049 GCTGTCTGGGGTAGCTGGAGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAAG 1108
Db 361 GCTGTCTGGGGTAGCTGGAGCACTTACTCAGGTGGCTGGTGAATGACACCTCAGAAG 420
QY 1109 GAATGAGTGTATAGAGA 1126
Db 421 GAATGAGTGTATAGAGA 438

RESULT 13
G25553 464 bp DNA linear STS 02-JUN-1996
DEFINITION human STS EST47735, sequence tagged site.
ACCESSION G25553
VERSION G25553.1 GI:1347785
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Hudson,T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically
JOURNAL Mapped STS
COMMENT Unpublished (1995)
Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: GATCGCCAGGCTGATTGAG
Primer B: CACCAAAACCTGGAAAATCA
STS size: 150
PCR Profile:
Preboak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Tag Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM
Tris-HCL: 10 mM
pH: 9.3
```


Search completed: July 4, 2005, 01:46:41
Job time : 5116.25 secs

```
RESULT 15
AX887897      AX887897      500 bp      DNA      linear      PAT 18-DEC-2003
LOCUS
DEFINITION   Sequence 3760 from Patent EP1033401.
ACCESSION   AX887897
VERSION     AX887897.1  GI:40046517
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
AUTHORS     Dunas Milne Edwards J.B., Duclert A. and Giordano J.Y.
TITLE       Expressed sequence tags and encoded human proteins
JOURNAL     Patent: EP 1033401-A 3760 06-SEP-2000;
            Genset (FR)
FEATURES
            Location/Qualifiers
            source          1..500
                             /organism="Homo sapiens"
                             /mol_type="unassigned DNA"
                             /db_xref="taxon:9606"
            CDS             70..381
                             /note="unnamed protein product"
                             /codon_start=1
                             /protein_id="CAE93822.1"
                             /db_xref="GI:40046518"
                             /translation="MAAAVPRAAFSLPLLGLLFLLSAPHGGSLHTKGALPLDTV
                             TFYKVPKSKFVLKFDTPYGEKQDEFKRLAENSASSDDLLVAEVGISDYGDRTW
                             S"
            sig_peptide     70..129
                             /note="score 6.9 seq AAFLSPLPLLGL/FL"
ORIGIN
Query Match      33.4%; Score 402.2; DB 6; Length 500;
Best Local Similarity 96.8%; Pred. No. 1.1e-102;
Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;
QY      26  CACGTGACCGCTGACTCGGGGGTTCTCCACTATCGCTTACCTACCTCCCTCTCGCAGGAA 85
DB      1  CACGTGACCGCTGACTCGGGGGTTCTCCACTATCGCTTACCTACCTCCCTCTCGCAGGAA 60
QY      86  CCCGCGATATGGCTGCGCTGTGCCCGCGCGGATTTCTCCCGCGCTT-CCTT 144
DB      61  CCCGCGATATGGCTGCGCTGTGCCCGCGCGGATTTCTCTCCCGCTTCCCTT 120
QY      145  CTCCTGGGCTTCTCTCTCTCTCTCGCTCCGATCGCGCGGCGCTGCACACAA-GGC 203
DB      121  CTCCTGGGCTTCTCTCTCTCTCTCGCTCCGATGCGCGGCGCTGCACACAAAGGCGC 180
QY      204  GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCAATCCCAAGCAAGTTGTC-TG 262
DB      181  GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCAATCCCAAGCAAGTTGTCCTTG 240
QY      263  GTGAAGTTTCGACACCCAGTACCCCTACGCTGAGAGCAGGATGAGTTCAAGCGTCTT-CT 321
DB      241  GTGAAGTTTCGACACCCAGTACCCCTACGCTGAGAGCAGGATGAGTTCAAGCGTCTTGTCT 300
QY      322  GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 381
DB      301  GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 360
QY      382  ACAAGCTGAAATGAGCTGAGTGAGAAATACAGCTGGAACAAGAGAGCTACCCA-TCCT 440
DB      361  GACAAGCGAATGAGCTGAGTGAGAAATACAGCTGGAACAAGAGAGCTACCCAGTCT 420
QY      441  TCTACTCTTCGGGATGGGACTTTGAGACCCAGTCCCATACACTGGGGCAGTT--AG 498
DB      421  TCTACTCTTCGGGATGGGACTTTGAGAA-CCAGTCCCATACACTGGGGCAGTTAAAG 479
QY      499  GTTGAGCCATCCAGCGCTGG 519
DB      480  GTTGAGCCATCCAGCGCTGG 500
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 615.67 Seconds
(without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-7
Perfect score: 1205
Sequence: 1 GCGCGCGAGTCCGCCCCC.....TCAGTGTTCGTTAAGTGG 1205

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq 16Dec04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1199.4	99.5	1205	AAV09021	Homo sapi
2	1062.8	88.2	1334	AACT78172	Human can
3	980	81.3	1600	AAI59481	Human pol
4	857.4	71.2	1242	AAI78847	DNA encod
5	803.8	66.7	963	ADB75219	Prostate
6	803.8	66.7	963	ACN39561	Tumour-as
7	638.4	53.0	1018	AAI83795	DNA encod
8	517.2	42.9	4529	ADBS8829	Toxicity-
9	517.2	42.9	4529	ADBS3548	Primary r
10	517.2	42.9	4529	ABT42350	Toxicity
11	496.6	41.2	1139	ADP72860	Renal tox
12	403	33.4	464	AAT22418	Human gen
13	402.2	33.4	500	AAO33762	Human sec
14	369	30.6	2087	AAI83797	DNA encod
15	328.8	27.3	395	ADT95875	Colon can
16	321	26.6	405	ACH18638	Human adu
17	312.8	26.0	629	ADQ051247	Novel can
18	279.2	23.2	450	AAI12456	Probe #23
19	279.2	23.2	450	ABA54164	Human foe
20	279.2	23.2	450	AAI33810	Probe #24

21	279.2	23.2	450	4	ABA43707	Human bre
22	279.2	23.2	450	4	ABA23909	Probe #23
23	279.2	23.2	450	4	AAK27875	Human bon
24	279.2	23.2	450	4	AAK02434	Human bra
25	279.2	23.2	450	4	ABS27459	Human liv
26	279.2	23.2	450	5	AAI02367	Probe #23
27	279.2	23.2	450	6	ABS02335	Human gen
28	279.2	23.2	450	9	ACH31743	Human bon
29	269.6	22.4	542	5	AAI78846	DNA encod
30	220.2	18.3	360	6	AAI45000	Psaomomys
31	213.8	17.7	383	8	ABX49602	Bovine ES
32	191.2	15.9	363	3	AAI04224	Human sec
33	190	15.8	343	4	AAI61267	Human pol
34	165	13.7	165	11	ADT97347	Colon can
35	159.8	13.3	163	5	AAI02030	Bladder c
36	159.8	13.3	163	10	ABX56173	Human DNA
37	151.4	12.6	415	8	ABX47558	Bovine ES
38	119.6	9.9	697	5	AAI83794	DNA encod
39	110.8	9.2	487	10	ACD93429	Human col
40	99.6	8.3	119	4	AAI21649	Probe #11
41	99.6	8.3	119	4	ABA66731	Human foe
42	99.6	8.3	119	4	AAI46937	Probe #15
43	99.6	8.3	119	4	ABA48813	Human bre
44	99.6	8.3	119	4	ABA33789	Probe #12
45	99.6	8.3	119	4	AAK40882	Human bon

ALIGNMENTS

RESULT 1
AAV09021
ID AAV09021 standard; cDNA; 1205 BP.
XX AC AAV09021;
XX DT 21-JUL-1998 (first entry)
XX DE Homo sapiens 20ql3 amplicon lb4 transcript.
XX KW 20ql3 amplicon; chromosome 20; tumour; detection;
XX KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
XX KW treatment; age-related macular degeneration; retinitis pigmentata;
XX KW Leber's congenital amaurosis; serine threonine kinase; db.
XX OS Homo sapiens.
XX PN WO9802539-A1.
XX PD 22-JAN-1998.
XX PF 15-JUL-1997; 97WO-US012343.
XX PR 15-JUL-1996; 96US-00680395.
XX PR 16-OCT-1996; 96US-00731499.
XX PA 17-JAN-1997; 97US-00785532.
XX PA (REGC) UNIV CALIFORNIA.
XX PI Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX DR New sequences from the 20ql3 amplicon - used for detecting chromosomal
XX PT abnormalities, particularly tumours, and for developing products for
XX PT treating diseases.
XX PS Claim 1; Page 63; 91pp; English.
XX CC The sequence is that of a cDNA sequence lb4 which was isolated from the
XX CC 20ql3 amplicon. It encodes a serine threonine kinase and can be used as a
XX CC probe for the detection of chromosomal abnormalities at 20ql3. It and
XX CC other sequences isolated from the 20ql3 amplicon are consistently

CC AAB43398 to AAB44239. The proteins can have activities based on the
 CC tissues and cells the genes are expressed in. Example of activities
 CC include: cytostatic; proliferative; antineoplastic; immunomodulator;
 CC antidiabetic; antiaesthetic; antirheumatic; antiarthritic;
 CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
 CC dermatologic; neuroprotective; cardiatic; thrombolytic; coagulant;
 CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
 CC polynucleotides and polypeptides can be used for preventing, treating or
 CC ameliorating medical conditions and diagnosing pathological conditions.
 CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
 CC the present invention may be used to treat immune disorders by activating
 CC or inhibiting the proliferation, differentiation or mobilisation of
 CC immune cells, to treat disorders of haematopoietic cells, autoimmune
 CC disorders, allergic reactions, graft versus host disease and organ
 CC rejection, modulate haemostatic or thrombolytic activity, modulate
 CC inflammation, cancers, cardiovascular disorders, neurological disease and
 CC bacterial or viral infections. The peptides, nucleotides, antibodies,
 CC agonists and antagonists may be also be used in drug screens. AAC78449 to
 CC AAC78457 and AAB44240 represent sequences used in the exemplification of
 CC the present invention

XX SQ Sequence 1334 BP; 337 A; 319 C; 396 G; 278 T; 0 U; 4 Other;

Query Match 88.2%; Score 1062.8; DB 3; Length 1334;

Best Local Similarity 98.2%; Pred. No. 6.5e-307;

Matches 1193; Conservative 5; Mismatches 5; Indels 12; Gaps 12;

QY 1 GCGCGGTGAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGCGTCTTCCATCATC 60
 DB 40 GCGCGGTGAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGCGTCTTCCATCATC 99
 QY 61 GCTTACCTACCTCCCTCTGAGGAACCCGGCGATATGGCTGCGCTGCGCCCGCGCGC 120
 DB 100 GCTTACCTACCTCCCTCTGAGGAACCCGGCGATATGGCTGCGCTGCGCCCGCGCGC 159
 QY 121 ATTCTCTCCCGCTGCTT-CCCTCTCTCCGGCTTCTGCTCTCTCCGCTCGCATGG 179
 DB 160 ATTCTCTCCCGCTGCTTCCCTCTCTCCGGCTTCTGCTCTCTCCGCTCGCATGG 219
 QY 180 CGCAGCGCGCTGCACACAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGT 238
 DB 220 CGCAGCGCGCTGCACACAAAGGCGCCCTTCCCTGGATACGGTCACTTTCTACAAGGT 279
 QY 239 CATTCCTCAAGAGATTCGTC-TGGTGAAGTTGACACCCAGTACCCCTAGCGTGAGAA 297
 DB 280 CATTCCTCAAGAGATTCGTC-TGGTGAAGTTGACACCCAGTACCCCTAGCGTGAGAA 339
 QY 298 CGAGATGAGTTCAAGCGTCTT-CTGAAACTCGGCTTCCAGCGATGATCTTGTGGC 356
 DB 340 CGAGATGAGTTCAAGCGTCTTCTGAAACTCGGCTTCCAGCGATGATCTTGTGGC 399
 QY 357 AGAGTGGGGATCTCAGATTAT-GTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAA 415
 DB 400 AGAGTGGGGATCTCAGATTATGTTGACAGCTGAACATGGAGCTGAGTGAGAAATACAA 459
 QY 416 GCTGACAAAGAGAGCTACCA-TCTTCTACCTCTTCCGGATGGGACTTTGAGAACCC 474
 DB 460 GCTGACAAAGAGAGCTACCACTCTTACCTCTTCCGGATGGGACTTTGAGAACCC 519
 QY 475 AGTCCATACACTGGGCGAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGAGGGGCAAGG 533
 DB 520 AGTCCATACACTGGGCGAGTTAAGTTGGAGCCATCCAGCGCTGGCTGAGGGGCAAGG 579
 QY 534 GGTCTACCTAGTATGCTGG-TGCTGCTGTATACGACGCCCTTGGCGGGAGTTTCAT 592
 DB 580 GGTCTACCTAGTATGCTGGTGTGCTGTATACGACGCCCTTGGCGGGAGTTTCAT 639
 QY 593 CAGGCGCTCTGTTGGAGG-CCGCGAGCCCTCTTGAAGCAGGGGCGGCAAGATAACCTCTC 651
 DB 640 CAGGCGCTCTGTTGGAGGCGCGCCCTCTTGAAGCAGGGGCGGCAAGATAACCTCTC 699
 QY 652 AAGTGTGAAGGAGCTCAGAAG-AGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGAT 710

DB 700 AAGTGTGAAGGAGACTCAGRAAGTGGCCGCGCAATACCTGAAGATCATGGGGAAGAT 759
 QY 711 CTTAGACAAGGGGAGACATT-CCAGATCAGAGATGACACGGATCGCCAGGCTGATTGA 769
 DB 760 CTTAGACAAGGGGAGGACTTTCCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGA 819
 QY 770 GAACAACAAGATGATGAC-GSCAAGAGGAGGAGCTCCAGAAAGAGCTTTAAACATCCTGAC 828
 DB 820 GAACAACAAGATGATGACGCGAAGAGGAGGAGCTCCAGAAAGAGCTTTAAACATCCTGAC 879
 QY 829 TGCCTTCCAGAAGAGGGGCGCGAAGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTC 888
 DB 880 TGCCTTCCAGAAGAGGGGCGCGAAGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTC 939
 QY 889 CAGGTTTGGTGGGGTAGGGAGGGAGTAACTGCTGCTGGCTGTGANTCCCTTGTGGA 948
 DB 940 CAGGTTTGGTGGGGTAGGGAGGGAGTAACTGCTGCTGGCTGTGANTCCCTTGTGGA 999
 QY 949 ATATAAGGGGYSKSGGAAAGGAGTAACTGCTGCTGAGCTTCTGAGCCCTGAGTATGCT 1008
 DB 1000 ATATAAGGGGAGTGGGAAAGTGGTAACTGCTGAGCTTCTGAGCCCTGAGTATGCT 1059
 QY 1009 GGACATTGATCTAACATGACCATGCTTGGGATGTCTTAGCTGCTTGGGGATAGCTGG 1068
 DB 1060 GGACATTGATCTAACATGACCATGCTTGGGATGTCTTAGCTGCTTGGGGATAGCTGG 1119
 QY 1069 AGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAAGAGTGAAGTCTATAGAGAGG 1128
 DB 1120 AGCACTTACTCAGTGGCTGGTGAATGACACCTCAGAAAGAGTGAAGTCTATAGAGAGG 1179
 QY 1129 AGAGAGGAGTACTGCCAGGCTTTTGACAGATGTAATCTTCATTCATTAAGTTTCA 1188
 DB 1180 AGAGAGGAGTACTGCCAGGCTTTTGACAGATGTAATCTTCATTCATTAAGTTTCA 1239
 QY 1189 GTGTTTGGTTAAGT 1203
 DB 1240 GTGTTTGGTTAANT 1254

RESULT 3

AAI59481

ID AAI59481 standard; cDNA; 1600 BP.

XX AAI59481;

DT 22-OCT-2001 (first entry)

XX Human polynucleotide SEQ ID NO 1684.

XX Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.

XX Homo sapiens.

XX WO200153312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000WO-US034263.

XX 23-DEC-1999; 99US-00471275.

XX 21-JAN-2000; 2000US-00488725.

XX 25-APR-2000; 2000US-0052317.

XX 20-JUN-2000; 2000US-00598042.

XX 19-JUL-2000; 2000US-00620312.

XX 03-AUG-2000; 2000US-00653450.

XX 14-SEP-2000; 2000US-00662191.

XX 19-OCT-2000; 2000US-00693036.

XX 29-NOV-2000; 2000US-00727344.

XX	(HYSE-) HYSEQ INC.	
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;	
XX	PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;	
XX	PI Zhou P, Goodrich R, Drmanac RT;	
XX	XX WPI; 2001-442253/47.	
XX	DR P-PSDB; AAM40325.	
XX	PT Novel nucleic acids and polypeptides, useful for treating disorders such	
XX	PT as central nervous system injuries.	
XX	PS Claim 1; SEQ ID NO 1684; 10078pp; English.	
XX	XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the	
XX	CC encoded polypeptides (AAM38642-AAM42213) with nootropic,	
XX	CC immunosuppressant and cytostatic activity. The polynucleotides are useful	
XX	CC in gene therapy. A composition containing a polypeptide or polynucleotide	
XX	CC of the invention may be used to treat diseases of the peripheral nervous	
XX	CC system, such as peripheral nervous injuries, peripheral neuropathy and	
XX	CC localised neuropathies and central nervous system diseases, such as	
XX	CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic	
XX	CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the	
XX	CC utilisation of the activities such as: Immune system suppression,	
XX	CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic	
XX	CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,	
XX	CC assays for receptor activity, arthritis and inflammation, leukaemias and	
XX	CC C.N.S disorders. Note: The sequence data for this patent did not form	
XX	CC part of the printed specification	
XX	XX Sequence 1600 BP; 402 A; 353 G; 431 G; 414 T; 0 U; 0 Other;	
XX	XX Query Match 81.3%; Score 980; DB 4; Length 1600;	
XX	XX Best Local Similarity 97.9%; Pred. No. 4.3e-282;	
XX	XX Matches 1112; Conservative 5; Mismatches 7; Indels 12; Gaps 12;	
Qy	82 GGAACCCCGCGATATGGCTGCGGTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-C 140	
Db	80 GAATTCGGCGATATGGCTGCGGTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTC 139	
Qy	141 CTTCTCTCGGGCTTCCTGCTCTCTCGCTCCGATGGCGGCGCGCTGCGACACCAA 200	
Db	140 CTTCTCTCGGGCTTCCTGCTCTCTCGCTCCGATGGCGGCGCGCTGCGACACCAA 199	
Qy	201 -GGCGCCCTCCCTGGATAGCTGCTCTTCTAAGGTGCTATCCCAAGCAAGTTGCT 259	
Db	200 GGGCGCCCTCCCTGGATAGCTGCTCTTCTAAGGTGCTATCCCAAGCAAGTTGCT 259	
Qy	260 C-TGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAGGAGGATGAGTTCAAGCGTCT 318	
Db	260 CTGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAGGAGGATGAGTTCAAGCGTCT 319	
Qy	319 T-CTGAAACCTCGGCTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTA 377	
Db	320 TGCTGAAACCTCGGCTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTA 379	
Qy	378 T-GTGACAGCTGACATGAGAGCTGAGTGAGAAATACAACTGCGACAAAGAGAGCTACCC 436	
Db	380 TGGTGACAGCTGACATGAGAGCTGAGTGAGAAATACAACTGCGACAAAGAGAGCTACCC 439	
Qy	437 A-TCTTCTACCTCTTCCGGGATGGGAGCTTTGAGAAACCCAGTCCCATACACTGGGCGAGT 495	
Db	440 AGTCTTCTACCTCTTCCGGGATGGGAGCTTTGAGAAACCCAGTCCCATACACTGGGCGAGT 499	
Qy	496 T-AGGTTGAGCCATTCACGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTCTGG 554	
Db	500 TAAGGTTGAGCCATTCACGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTCTGG 559	
Qy	555 -TGCTGCTGTATACGACGCCCTGGCGCGGGAGTTTCATCAGGGCTCTGCTGTTGGAGG- 612	
Db	560 TTGCTGCTGTATACGACGCCCTGGCGCGGGAGTTTCATCAGGGCTCTGCTGTTGGAGGC 619	

Qy	613 CCGCCAGGCGCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAA 672	
Db	620 CCGCCAGGCGCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAA 679	
Qy	673 G-AGTGGGCGAGCAATACCTGAAGATCATGGGAGAGATCTTAGACCAAGGGGAGCACTT 731	
Db	680 GAAAGTGGGCGAGCAATACCTGAAGATCATGGGAGAGATCTTAGACCAAGGGGAGCACTT 739	
Qy	732 -CCAGCATCAGAGATCACCGGATCGCCAGGCTGATTGAGAAGAAACAAGATGAGTGAC-G 789	
Db	740 CCAGCATCAGAGATCACCGGATCGCCAGGCTGATTGAGAAGAAACAAGATGAGTGACGG 799	
Qy	790 GCAGAAGGAGAGCTCCAGAAGAGCTTAAACATCCTCACTGCTCCCTCCAGAAGAGGGGGC 849	
Db	800 GAAGAAGGAGAGCTCCAGAAGAGCTTAAACATCCTCACTGCTCCCTCCAGAAGAGGGGGC 859	
Qy	850 CGAGAAGAGAGAGCTGTAAAAGAGGCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGG 909	
Db	860 CGAGAAGAGAGAGCTGTAAAAGAGGCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGG 919	
Qy	910 AGGGGANAAGTTAACTGCTGCTGTGANTCCCTTGTGGAATATAAGGGGGMKGGGAAA 969	
Db	920 AGGGGAGAGTTAACTGCTGCTGTGAGTCCCTTGTGGAATATAAGGGGGMKGGGAAA 979	
Qy	970 AGMGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTTGACATTCATGCTAACATGAC 1029	
Db	980 AGTGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTTGACATTCATGCTAACATGAC 1039	
Qy	1030 CATGCTTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGG 1089	
Db	1040 CATGCTTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGG 1099	
Qy	1090 TGAATGACACCTCAGAAAGGAATGAGTGTCTATAGAGAGGAGAGAGTGTACTGCCCCAG 1149	
Db	1100 TGAATGACACCTCAGAAAGGAATGAGTGTCTATAGAGAGGAGAGAGTGTACTGCCCCAG 1159	
Qy	1150 GTCTTTTCAGAGATTAATTCATTCAATTAAGTTTCAAGTTTTCAGTTTTCAGTTTTCAGTTT 1205	
Db	1160 GTCTTTTCAGAGATTAATTCATTCAATTAAGTTTTCAGTTTTCAGTTTTCAGTTTTCAGTTT 1215	

RESULT 4

AAS78847

ID AAS78847 standard; cDNA; 1242 BP.

XX AC AAS78847;

XX AC AAS78847;

XX DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #14651.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

XX WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSE-) HYSEQ INC.

XX PA Drmanac RT, Liu C, Tang YT;

XX PI WPI; 2001-639362/73.

XX DR P-PSDB; ABG14660.

XX PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.

Claim 1; SEQ ID NO 14651; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published pct sequences

Sequence 1242 BP; 290 A; 300 C; 378 G; 274 T; 0 U; 0 Other;

Query Match	71.2%;	Score 857.4;	DB 5;	Length 1242;
Best Local Similarity	93.7%;	Pred. No. 1.8e-245;		
Matches 1163;	Conservative 5;	Mismatches 33;	Indels 40;	Gaps 25;
QY	1	GC	GC	GC
QY	1	GC	GC	GC
QY	61	GC	GC	GC
QY	61	GC	GC	GC
QY	121	AT	AT	AT
QY	121	AT	AT	AT
QY	180	CG	CG	CG
QY	181	CG	CG	CG
QY	238	TC	TC	TC
QY	241	TC	TC	TC
QY	297	AG	AG	AG
QY	301	AG	AG	AG
QY	354	GC	GC	GC
QY	361	GC	GC	GC
QY	411	TACA	TACA	TACA
QY	421	TACA	TACA	TACA
QY	467	-GAGAA	-GAGAA	-GAGAA
QY	481	GAGG	GAGG	GAGG
QY	520	CTGAA	CTGAA	CTGAA
QY	541	CTGAA	CTGAA	CTGAA

XX Schlegel R, Monahan JE, Endege WO, Gannavarapu M, Gorbacheva B;
PI Hoerish S, Kamathkar S, Wonsey AM, Glatt K, Zhao X, Anderson D;
XX WPI; 2003-248033/24.
XX New nucleic acid molecule, useful for diagnosing or treating prostate
PT cancer.
XX Disclosure; SEQ ID NO 43; 99pp; English.
XX The invention relates to newly discovered cancer markers associated with
CC the cancerous state of prostate cells. Also disclosed is a method of
CC assessing whether a patient is afflicted with prostate cancer. The method
CC of the invention involves assessing whether a patient is afflicted with
CC prostate cancer by comparing the level of expression of a marker in a
CC patient sample and the normal level of expression of the marker in a
CC control non-prostate cancer sample, where a significant increase in the
CC level of expression of the marker in the patient sample and the normal
CC level indicates that the patient is afflicted with prostate cancer.
CC Nucleic acids of the invention are useful for diagnosing or treating
CC prostate cancer, and may be useful in gene therapy. Sequences given in
CC ADB7517-ADB75631 represent marker cDNA and proteins. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 963 BP; 225 A; 240 C; 297 G; 201 T; 0 U; 0 Other;

Query Match 66.7%; Score 803.8; DB 10; Length 963;
Best Local Similarity 97.8%; Pred. No. 1.8e-229;
Matches 934; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 87 CCGCGGATATGGCTGCGCGCTGTGCCCCGCGCGCATTTCTCCCGCGTGT-CCCTTC 145
Db 4 CCGCGGATATGGCTGCGCGCTGTGCCCCGCGCGCATTTCTCCCGCGTGT-CCCTTC 63
Qy 146 TCCTTGGGCTTCTGCTCTCTCCGCTCGCATGGCGGCGGCTGCACACCAA-GGCG 204
Db 64 TCCTTGGGCTTCTGCTCTCTCCGCTCGCATGGCGGCGGCTGCACACCAA-GGCG 123
Qy 205 CCCTTCCCTCGATACCGGTCACTTTCTACAAAGGTCAITTCACAAAGCAAGTTCGTG 263
Db 124 CCCTTCCCTCGATACCGGTCACTTTCTACAAAGGTCAITTCACAAAGCAAGTTCGTG 183
Qy 264 TGAAGTTGACACCCAGTACCCCTACCGTGAGAGCAGGATGAGTTCAGGCTCT-CTG 322
Db 184 TGAAGTTGACACCCAGTACCCCTACCGTGAGAGCAGGATGAGTTCAGGCTCTGCTG 243
Qy 323 AAAAATCTCGGCTTCCAGCGATGATCTTTGGTGAGAGGTTGGGATCTCAGATTAT-GTG 381
Db 244 AAAATCTCGGCTTCCAGCGATGATCTTTGGTGAGAGGTTGGGATCTCAGATTATGTTG 303
Qy 382 ACAAGCTGAAATGAGAGTGTGAGAAATCAAGCTGGAACAAAGAGCTACCGCA-TCT 440
Db 304 ACAAGCTGAAATGAGAGTGTGAGAAATCAAGCTGGAACAAAGAGCTACCGCAAGTCT 363
Qy 441 TCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGG 499
Db 364 TCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTTAAGG 423
Qy 500 TTGAGAGCATCCAGGCTGGCTGAAGGGCGAAGGGGTCTACTAGGTATGCTGTG-TGCC 558
Db 424 TTGAGAGCATCCAGGCTGGCTGAAGGGCGAAGGGGTCTACTAGGTATGCTGTGTGCC 483
Qy 559 TGCTGTATACAGCCCTGCGCGGGGAGTTTCATCAGGGCTCTGTGTGTGAGG-CCGCC 617
Db 484 TGCTGTATACAGCCCTGCGCGGGGAGTTTCATCAGGGCTCTGTGTGTGAGGCCCCGCC 543
Qy 618 AGGCCCTTTGAAGCGGGCAAGATACCTCTCAAGTGTGAGAGACTCAGAG-AGT 676
Db 544 AGGCCCTTTGAAGCGGGCAAGATACCTCTCAAGTGTGAGAGACTCAGAGAAAGT 603

Qy 677 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGAGCACATT-CCAG 735
Db 604 GGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGAGGACTTCCAG 663
Qy 736 CATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAGAACAGATGAGTGACCGC-AGA 794
Db 664 CATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAGAACAGATGAGTGACCGCAAGA 723
Qy 795 AGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAAAGAGGGGCCGAGA 854
Db 724 AGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAAAGAGGGGCCGAGA 783
Qy 855 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGAGGGG 914
Db 784 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGAGGGG 843
Qy 915 ANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYNKSGGAAAGWGG 974
Db 844 AGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGGTGTAGTGGAAACAGTGG 903
Qy 975 TACTAACCCACGATTCTGAGCCCTGAGTATGCTGGACATTGATGCTAAACATGAC 1029
Db 904 TACTAACCCACGATTCTGAGCCCTGAGTATGCTGGACATTGATGCTAAACATGAC 958
RESULT 6
ACN39561
ID ACN39561 standard; cDNA; 963 BP.
XX ACN39561;
XX ACN39561;
DT 18-NOV-2004 (first entry)
XX Tumour-associated antigenic target (TAT) cDNA DNA325764, SEQ ID NO:3785.
DE Tumour-associated antigenic target; TAT; human; overexpression; cancer;
XX tumour; diagnosis; cell proliferative disorder; breast cancer;
XX colorectal cancer; lung cancer; ovarian cancer; liver cancer;
XX central nervous system cancer; bladder cancer; pancreatic cancer;
XX cervical cancer; melanoma; leukaemia; hybridisation probe;
XX chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic; gene; ss.
XX Homo sapiens.
XX WO2004030615-A2.
XX 15-APR-2004.
XX 29-SEP-2003; 2003WO-US028547.
XX 02-OCT-2002; 2002US-0414971P.
XX (GETH) GENENTECH INC.
XX Wu TD, Zhang Z, Zhou Y;
XX WPI; 2004-347921/32.
XX P-PSDB; ABM81468.
XX New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX Claim 1; SEQ ID NO 3785; 7273pp; English.
XX The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and

SQL	Sequence	1018 BP; 221 A; 257 C; 310 G; 230 T; 0 U; 0 Other;
	Query Match	53.0%; Score 638.4; DB 5; Length 1018;
	Best Local Similarity	80.5%; Pred. No. 5.7e-180;
	Matches	972; Conservative 5; Mismatches 8; Indels 223; Gaps 11;
Qy	1	CGCGCGTGAAGTCGCGCCCCCGACGTCAGCTGACCGCTGACTCGGGGGTTCTCCACTATC 60
Db	25	CGCGCGTGAAGTCGCGCCCCCGACGTCAGCTGACCGCTGACTCGGGGGTTCTCCACTATC 84
Qy	61	GCTTACTACTCTCCCTCTGCAAGAACCGCGGATATGGCTGCCCTGTCGCCCGCGCGC 120
Db	85	GCTTACTACTCTCCCTCTGCAAGAACCGCGGATATGGCTGCCCTGTCGCCCGCGCGC 144
Qy	121	ATTTCTCTCCCGCTGCTT-CCCTTCTCTGGGCTTCTCTGCTCTCTCCGCTCCGATGG 179
Db	145	ATTTCTCTCCCGCTGCTTCCCTTCTCTGGGCTTCTCTGCTCTCTCCGCTCCGATGG 204
Qy	180	CGGAGCGGCTGCACACCAA-GCGGCCCTTCCCTGGATACGCTCACTTTCTACAAGT 238
Db	205	CGGAGCGGCTGCACACCAAAGGCGCCCTTCCCTGGATACGCTCACTTTCTACAAGT 264
Qy	239	CATTCCCAAGCAAGTTCGTCTGTGAGTTCCACACCCAGTACCCCTACGGTGAGAAG 298
Db	265	TAT----- 267
Qy	299	CAGGATGAGTTCAAGCGTCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGTTGGCAG 358
Db	268	----- 267
Qy	359	AGTGGGGATCTCAGATTATGTGACAGCTGAAATGAGTGAAGTGAAGTAACAAGCT 418
Db	268	-----GGTGACAGCTGAACATGGAGCTGAGTGAGAAATACAAGCT 308
Qy	419	GGACAAAGAGCTACCCA-TCCTTCTACCTCTTCCGGATGGGACCTTGAGAACCCAGT 477
Db	309	GGACAAAGAGCTACCCAGCTTCTACCTCTTCCGGATGGGACCTTGAGAACCCAGT 368
Qy	478	CCCATACACTGGGCGAGTT-AGGTTGGAGCATCCAGCGCTGGCTGAAGGGGCAAGGGT 536
Db	369	CCCATACACTGGGCGAGTTAGGTTGGAGCATCCAGCGCTGGCTGAAGGGGCAAGGGT 428
Qy	537	C-TACCTAGGATGCTCTGG-TGCTGCTGTGTATACGACGCCCTCGCGGGGAGTTCACTCA 594
Db	429	CTTACCTAGGATGCTCTGGTTGCTGTGTATACGACGCCCTCGCGGGGAGTTCACTCA 488
Qy	595	GGGCTCTGTGTCTGGAGG-CGCGACGGCCCTTGAAGCAGGCGGCAAGATTAACCTCTCAA 653
Db	489	GGGCTCTGTGTCTGGAGGCGCGCGGCGGCGGCTTGAAGCAGGCGGCAAGATTAACCTCTCAA 548
Qy	654	GTGTGAAGGAGCTCAGAG-AGTGGGCGGAGCAATACCTCAAGATCATGGGGAAGATCT 712
Db	549	GTGTGAAGGAGCTCAGAGAGTGGGCGGAGCAATACCTCAAGATCATGGGGAAGATCT 608
Qy	713	TAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACACGGATCCCGAGGCTGATGAGA 771
Db	609	TAGACCAAGGGGAGCACTTCCAGCATCAGAGATGACACGGATCCCGAGGCTGATGAGA 668
Qy	772	AGACACAGATGATGACGCGAGAGAGGAGGAGCTCCAGNAGAGCTTAACATCTCTGACTGC 831
Db	669	AGAAC----- 673
Qy	832	CTTCCAGAAGAGGGGCGGAGAGGAGCTGTAAAAAGGCTGTGTGATTTTCAG 891
Db	674	-----HAGAGAGCTGTAAAAAGGCTGTGTGATTTTCGG 710
Qy	892	GGTTTGTGGGGTAGGGAGGGAGNAGTAACTGCTGGCTGTGANTCCCTTTGGGAATA 951
Db	711	GGTTTGTGGGGTAGGGAGGGAGAGTAACTGCTGGCTGTGAGTCCCTTTGGGAATA 770
Qy	952	TAAAGGGGYSKGGGAAAGHGTACTTAACCCAGATTCAGCCCTGAGTATCCCTGGA 1011
Db	771	TAAAGGGGAGTGGGAAAGTGGTACTTAACCCAGATTCAGCCCTGAGTATCCCTGGA 830

Qy	1012	CATTGATGCTAATCATGACCATGCTTGGGATGTCTCTAGCTGGTCTGGGGATAGCTGGAGC 1071
Db	831	CATTGATGCTAATCATGACCATGCTTGGGATGTCTCTAGCTGGTCTGGGGATAGCTGGAGC 890
Qy	1072	ACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGAATGAGTCTATAGAGAGAGA 1131
Db	891	ACTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGAATGAGTCTATAGAGAGAGA 950
Qy	1132	GAGGAGTGTACTGCCAGGTCTTTGACAGATGAATTTCTCATTCATCAATTAAGTTTCAGTG 1191
Db	951	GAGGAGTGTACTGCCAGGTCTTTGACAGATGAATTTCTCATTCATTAAGTTTCAGTG 1010
Qy	1192	TTTTGGTT 1199
Db	1011	TTTTGGTT 1018
RESULT 8		
ADBS8829	ID	ADBS8829 standard; DNA; 4529 BP.
XX	AC	ADBS8829;
XX	XX	
DT	DT	
XX	XX	04-DEC-2003 (first entry)
XX	DE	Toxicity-related gene, SEQ ID 3855.
XX	XX	Toxic; toxin; gene expression profile; hepatotoxicity; liver;
KW	KW	drug screening; toxicity assay; ds.
XX	OS	Unidentified.
PN	PN	WO2003064624-A2.
XX	XX	
PD	PD	07-AUG-2003.
XX	XX	
PF	PF	31-JAN-2003; 2003WO-US003194.
XX	XX	
PR	PR	31-JAN-2002; 2002US-00060087.
PR	PR	15-MAR-2002; 2002US-0364045P.
PR	PR	15-MAR-2002; 2002US-0364055P.
PR	PR	30-DEC-2002; 2002US-0436643P.
XX	PA	(GENE-) GENE LOGIC INC.
XX	PI	Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Elashoff M;
XX	XX	WPI; 2003-689530/65.
XX	XX	
PT	PT	Predicting a toxic effect of a compound, useful in identifying toxicity
PT	PT	markers in liver tissues or cells for drug screening and toxicity assays,
PT	PT	comprises preparing gene expression profile of tissue or cells exposed to
PT	PT	the compound.
XX	XX	Claim 1; SEQ ID NO 3855; 1156pp; English.
PS	PS	
XX	XX	The present invention relates to a method for predicting a toxic effect
CC	CC	of a compound. The method comprises preparing a gene expression profile
CC	CC	of a tissue or cell sample exposed to the compound, and comparing the
CC	CC	gene expression profile to a database comprising SEQ ID 1-4925, where
CC	CC	differential expression of the gene indicates at least one toxic effect.
CC	CC	The method is useful for predicting at least one toxic effect of a
CC	CC	compound, predicting hepatotoxicity or the progression of a toxic effect
CC	CC	of a compound, identifying an agent that modulates the onset or
CC	CC	progression of a toxic response, predicting the cellular pathways that a
CC	CC	compound modulates in a cell, and identifying an agent that modulates at
CC	CC	least one activity of a protein. The method and compositions of the
CC	CC	present invention using a database of genes having liver toxin-induced
CC	CC	differential expression, are useful in identifying toxicity markers in
CC	CC	liver tissues or cells for drug screening and toxicity assays. Note: The
CC	CC	sequence data for this patent did not form part of the printed
CC	CC	specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published/pct_sequences.

CC	at ftp.wipo.int/pub/published_pct_sequences.
XX	
SQ	Sequence 4529 BP; 1018 A; 1179 C; 1187 G; 1145 T; 0 U; 0 Other;
	Query Match 42.9%; Score 517.2; DB 10; Length 4529;
	Best Local Similarity 78.5%; Pred. No. 2.3e-143; Indels 26; Gaps 17;
	Matches 832; Conservative 1; Mismatches 201; Indels 26; Gaps 17;
Qy	1 GC GCGCGTGAGTCCGCCCCCGCCAGCTACGTCACCGCTGACTCGGGGGGTTCTCCACCTATC 60
Db	
Db	3277 GCGCGCGGGAGCGCCCTCCCACTGAGTGACGCCCACTCGAGCGCTC-CTGCTGTA 3335
Qy	61 GCTTACCTACCTCCCTCTGACAGAACCCGGCGATATGCTGCCGTGTGCCCGCGCGC 120
Db	
Db	3336 GCTTGCCTATCTCCCGCTCCACAAATCCGGCGTGATGGCTCCCGCGTTCCTGCGCGCGT 3395
Qy	121 ATTTCCTCCCGCTGCT-TCCCTTCTCCTGGGCTCTCTGCTCTCTCCGCTCCGCATGG 179
Db	
Db	3396 TTGCGCTCTCCCACTGCTGTCCGTTCTCTTGGCGTCTGCTCTCTCTGCTCTCATGG 3455
Qy	180 CGG CAGCGCGCTG CACACCAA -GGCGGCCCTTCCCTCGATACGGTCACTTCTTACAAAGT 238
Db	
Db	3456 CGC CAGCGGCTG CACACGAAGGCGCCCTCCCTTGGACACAGTCACTTCTTACAAAGT 3515
Qy	239 CATTTCCCAAAGCAAGTTTGTCT-TGTGAAAGTTCCGACACCCAGTACCCCTACCGTGA 297
Db	
Db	3516 CATTTCCCAAAGCAAGTTTGTCTTGTGTGAAGTTCCGACACCCAGTACCCCTATGGAGAA 3575
Qy	298 GCAGGATGAGTTCAAGCGTCT-TCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGC 356
Db	
Db	3576 GCAAGATGAGTTTAAAGCGTCTGGCTGAGAACTCAGCTCCAGCGATGATCTCTTGGTGGC 3635
Qy	357 AGAGTGGGGATCTCAGATTAT-CTGCAAGCTGAAATGAGAGCTGAGTGAGAAATCAA 415
Db	
Db	3636 AGAGTGGGGATCTCAGACTATGTGTACAGCTGAAATGAGAGCTGAGTGAGAGTACAA 3695
Qy	416 GCTGGACAAAGAGAGCTACCCA-TCTTCTACCTCTTCGGGATGGGACCTTTGAGAACCC 474
Db	
Db	3696 GCTGGACAAAGAGAGCTACCCAGTCTTCTACCTCTTCGGGATGGGACCTTTGAGAAATCC 3755
Qy	475 AGTCCCATACACTGGGCGAGTT-AGTTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 533
Db	
Db	3756 TGTCCCATACAGCGGGCGAGTTAAAGTTGGAGCCATCCAGCGCTGGCTCAAGGGGCGAGG 3815
Qy	534 GGTCTACTAGGTATGCTTG-TGCCTGCCCTGTATACGACGCCCTGGCGGGGAGTTTCAT 592
Db	
Db	3816 AGTCTATCTGGGCATGCTGAGATGTCTGCTGCTACATGCCCTGGCGGCGCATGTTTCAT 3875
Qy	593 CAGGGCTCTGTGTGGAGG-CCGCCAGGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTC 651
Db	
Db	3876 CGAGGCTCTCAGCAGAGAGGGCCCGCCAGGCCATCTTGAACAGGGGAGGATGGCCCTCTC 3935
Qy	652 AAGTGTGAAGGAGACTCAGAAAG-AGTGGCCGAGCAATATCCTGAAGATCATTGGGGAAGAT 710
Db	
Db	3936 AAGTGTGAAGGAGACAGACAAAGATGGGCGCAGTCAGTACCTGAAGATCATTGGGGAAGAT 3995
Qy	711 CTTAGACCAAGGGGAGCATT-CCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGA 769
Db	
Db	3996 CTTGGACCAAGTGAAGACTTCCCGGCTCCGAGCTGCCCGGATCAGTAAGCTCATTTGA 4055
Qy	770 GAAGAACCAAGATGATGACGCGAAGAGGAGGCTCCAGAAAGCTTTAAACATCTCGACT 829
Db	
Db	4056 GACCAAGATGAGTGAGG-GTAAAGAGGAGAGCTGCAGAGAGGCTCAACATCTCACC 4113
Qy	830 GCCTTCAGAAAGAGGGGCGGAGAAAGAGAGCTGTATAAAGCGTGTCTGTGATTTTCC 889
Db	
Db	4114 GCCTTCGCGAAGAAAGGCGCGAAGAGGAGAGCTGT-----GAGGGCACCGAGCTGCT 4167
Qy	890 AGGGTTTGTGGGGTATGGGAGGGGAGTATAAATCTGCTGGCTGTGANTCCCTTTGTGAA 949
Db	
Db	4168 AGGGTTTGTGAGGGCGGGGAGGG-AGAGCGCCCTCTGCTGACTGTGGGGCCCTGTGGG-- 4223
Qy	950 TATAAGGGGGYMSKGGGAAAAGWGCTACTAACCCACGATTTCTGAGCCCTGAGTATGCTGTG 1009

Db 4224 TGGAGGGGGCAGTGGAGCAAGCAGTCCTGAGCCAGAGACCTGTGCCCGAG--TGCCTG 4281
Qy 1010 GACATTGATCTCAACATGACCATGCTTGGGATGTCCTAG 1049
Db 4282 GACACTGAGGCTGCTGAGACCATAACCTTGACATCCTTGG 4321

RESULT 9
ADB53548
ID ADB53548 standard; DNA; 4529 BP.
XX AC ADB53548;
XX XX
XX 04-DEC-2003 (first entry)
XX XX
XX Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:4090.
XX toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX XX
OS Rattus norvegicus.
XX XX
XX WO2003065993-A2.
XX XX
XX 14-AUG-2003.
XX XX
XX 04-FEB-2003; 2003WO-US003482.
XX XX
XX 04-FEB-2002; 2002US-0353171P.
XX 13-MAR-2002; 2002US-0363534P.
XX 08-APR-2002; 2002US-0370248P.
XX 10-APR-2002; 2002US-0371134P.
XX 10-APR-2002; 2002US-0371135P.
XX 10-APR-2002; 2002US-0371150P.
XX 11-APR-2002; 2002US-03711413P.
XX 19-APR-2002; 2002US-0373601P.
XX 19-APR-2002; 2002US-0373602P.
XX 22-APR-2002; 2002US-0374139P.
XX 08-MAY-2002; 2002US-0378370P.
XX 09-MAY-2002; 2002US-0378652P.
XX 09-MAY-2002; 2002US-0378653P.
XX 09-MAY-2002; 2002US-0378665P.
XX 09-JUL-2002; 2002US-0394230P.
XX 09-JUL-2002; 2002US-0394253P.
XX 04-SEP-2002; 2002US-0407688P.
XX 28-JAN-2003; 2003US-0442900P.
XX XX
(GENE-) GENE LOGIC INC.
XX XX
XX Mendrick D, Porter M, Johnson K, Higgs B, Caastle A, Orr M;
XX Elashoff M;
XX XX
XX WPI; 2003-731472/69.
XX XX
XX Determining if a compound induces a toxic effect on a tissue or cell, for
PT identifying hepatotoxic compounds, comprises comparing a gene expression
PT profile of a tissue or cell sample to a database of Tox mean and non-Tox
PT mean values.
XX XX
XX Claim 44; SEQ ID NO 4090; 874pp; English.
XX XX
XX The present invention describes a method for determining whether a
CC compound induces a toxic effect on a tissue or cell. The method comprises
CC preparing a gene expression profile of a tissue or cell sample exposed to
CC the compound, and comparing the gene expression profile to a database
CC comprising data or information on the Tox mean and non-Tox mean value.
CC The method is useful for predicting or identifying at least one toxic
CC effect, particularly hepatotoxicity, of a test or unknown compound. The
CC genes listed in the specification are useful as diagnostic or toxicity
CC markers for the prediction or identification of the physiological state
CC of tissue or cell sample that has been exposed to a compound or to

Determining if a compound induces a toxic effect on a tissue or cell, for identifying hepatotoxic compounds, comprises comparing a gene expression profile of a tissue or cell sample to a database of Tox mean and non-Tox mean values.

Claim 44: SEO ID NO 4090: 874pp: English:

The present invention describes a method for determining whether a compound induces a toxic effect on a tissue or cell. The method comprises preparing a gene expression profile of a tissue or cell sample exposed to the compound, and comparing the gene expression profile to a database comprising data or information on the Tox mean and non-Tox mean value. The method is useful for predicting or identifying at least one toxic effect, particularly hepatotoxicity, of a test or unknown compound. The genes listed in the specification are useful as diagnostic or toxicity markers for the prediction or identification of the physiological state of tissue or cell sample that has been exposed to a compound or to

CC identify or predict the toxic effects of a compound or an agent. These
CC may also be used as markers for monitoring toxicity progression or for
CC drug screening. The present sequence represents a primary rat hepatocyte
CC toxicity modelling related gene sequence from the present invention.

XX	SQ	Sequence	4529 BP; 1018 A; 1179 C; 1187 G; 1145 T; 0 U; 0 Other;
Qy		Query Match	42.9%; Score 517.2; DB 10; Length 4529;
Db		Best Local Similarity	78.5%; Pred. No. 2.3e-143;
XX		Matches	832; Conservative 1; Mismatches 201; Indels 26; Gaps 17;
Qy	1	CGCGCGTGAGTCCGCCCCCAGTACAGTACCGCTGACTCGGGGGTTTCCACATATC 60	
Db	3277	CGCGCGGGGACGGCCCTCCAGTACGTCACCGCCCACTCGGAGCGCTC-CTGCTGTA 3335	
Qy	61	GCTTACCTACCTCCCTCTGCAGGAACCGCGGATATGCTGCGCTGTCGCCCGCGGC 120	
Db	3336	GCTTGCCCATCTCCCGCTCCCAATCCGGGATGATGCTGCCCGCTTCCTGGCGCGT 3395	
Qy	121	ATTTCTCTCCCGCTGCT-TCCCTTCTCTGGGCTTCTGCTCTCTCCGCTCCGCATGG 179	
Db	3396	TTGCTCTCTCCCACTGCTGTCGGTCTCTCTGGGCTGCTGCTCTCTGCTCTCATGG 3455	
Qy	180	CGGAGCGGCTGCACACCAA-GGGGCGCTTCCCTGTGATACGGTCACTTTCTACAAGT 238	
Db	3456	CGCCAGCGGCTGCACACGAAGGGCGCCCTTCCCTTGGACACAGTCACTTTCTACAAGT 3515	
Qy	239	CATTCCCAAAAGCAAGTTGCTC-TGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAGAA 297	
Db	3516	CATTCCCAAAAGCAAGTTGCTTGGTGAAGTTTCGACACCCAGTACCCCTATGAGAGAA 3575	
Qy	298	GCAGGATGAGTTCAAGCGTCT-TCTGAAACTCGGCTTCCACGATGATCTCTTGGTGGC 356	
Db	3576	GCAAGATGAGTTTAAAGCTCTGGCTGAGAACTCAGCCTCCAGCGATGATCTTGGTGGC 3635	
Qy	357	AGAGTGGGATCTCAGATTAT-GTGCAAGCTGAAATGAGCTGAGTGAGAAATACAA 415	
Db	3636	AGAGTGGGATCTCAGATATGTGTGCAAGCTGAACTGAGTGTGAGAAAGTACAA 3695	
Qy	416	GCTGCACAAAGAGAGCTACCCA-TCTTCTACTCTTCCGGATGGGACTTTGAGAACCC 474	
Db	3696	GCTGCACAAAGAGAGCTACCCAGTCTTCTACTCTTCCGGATGGGACTTTGAGAACCC 3755	
Qy	475	AGTCCATACACTGGGGCAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGG 533	
Db	3756	TGTCCTACAGCGGGGAGTTAAGTTGGAGCCATCCAGCGCTGGCTCAAGGGGCGAGG 3815	
Qy	534	GGTCTACTAGTATGCTCGG-TGCTGCTGTATACAGCCCTTGGCGGGGAGTTTCAT 592	
Db	3816	AGTCTATCTGGGCATGCTGTGATGCTGCTGCGTACGATGCCCTTGGCGGCGCAGTTTCAT 3875	
Qy	593	CAGGCGCTCTGTTGGAGG-CCGCCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTC 651	
Db	3876	CGAGCGCTCCAGCAGAGAGGGCCGCCAGGCCATCTGAAACAGGGGCGAGGATGGCTCTC 3935	
Qy	652	AAGTGTGAAGAGAGTCAAGAAG-AGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGAT 710	
Db	3936	AGGTTGAAGAGAGACAGACAGAGAGTGGGCCAGTCACTTGAAGATCATGGGGAAGAT 3995	
Qy	711	CTTAGACCAAGGGGAGCACTT-CCAGCATACAGATGACACGGATTCGCACGGCTGATTGA 769	
Db	3996	CTTGACCAAGTGAAGACTTCCCGGCTCCGAGCTGGCCCGGATCAGTAGCTCAITGA 4055	
Qy	770	GAAGACAGATGATGAGCGCAGAGGAGGCTCCAGAGAGCTTAAACATCTGACT 829	
Db	4056	GAACAAGATGATGAGG--GTAAGAGGAAGAGCTGCAGAGAGGCTCAACATCTCTACC 4113	
Qy	830	GCTTTCAGAAAGAGGGGCGGAGAAAGAGAGCTGTATAAAGGCTGTCTGTGATTTTCC 889	
Db	4114	GCTTTCGCAAGAAAGGCGCCGAGAGGAGGAGCTGT-----GAGGGCACCAGGCTGT 4167	
Qy	890	AGGTTTGGTGGGGTAGGGAGGGGAGGAGTTAACTGCTGCTGTGANTCCCTTGTGGAA 949	

Db	4168	AGGTTTGTCTAGGGCGGGGAGGG--AGAGCGCGCTCTGACTGTGTGGGGCCCTGTGGG-- 4223	
Qy	950	TATAAGGGGGMKGGGAAAGWGGTACTAAACCCACGATTTCTGAGCCCTGAGTATGCTG 1009	
Db	4224	TGGAGGGGCGAGTGGAGCAAGCAGTCTCTGAGCCAGAGACTGTGCCCCGAG--TGCTG 4281	
Qy	1010	GACATTCATCTAACATGACCATGCTTTGGGATGTTCTTAG 1049	
Db	4282	GACACTGAGGCTGCTGAGACCATACCTCGACATCTCTGG 4321	
XX	RESULT 10		
XX	ABT42350		
XX	ID	ABT42350 standard; DNA; 4529 BP.	
XX	AC	ABT42350;	
XX	DT	26-JUN-2003 (first entry)	
XX	XX	Toxicity modelling related rat gene SEQ ID No 2052.	
XX	DE	Toxic effect; gene expression profile; renal toxicity; toxicity marker;	
XX	KW	database; drug screening; toxicity assay; rat; ds.	
XX	OS	Rattus norvegicus.	
XX	PN	WO200295000-A2.	
XX	XX	28-NOV-2002.	
XX	PF	22-MAY-2002; 2002MO-US016173.	
XX	PR	22-MAY-2001; 2001US-0292335P.	
XX	PR	13-JUN-2001; 2001US-0297523P.	
XX	PR	19-JUN-2001; 2001US-0298925P.	
XX	PR	10-JUL-2001; 2001US-0303807P.	
XX	PR	10-JUL-2001; 2001US-0303808P.	
XX	PR	10-JUL-2001; 2001US-0303810P.	
XX	PR	28-AUG-2001; 2001US-0315047P.	
XX	PR	27-SEP-2001; 2001US-0324928P.	
XX	PR	22-OCT-2001; 2001US-0330462P.	
XX	PR	01-NOV-2001; 2001US-0330867P.	
XX	PR	21-NOV-2001; 2001US-0331805P.	
XX	PR	06-DEC-2001; 2001US-0336144P.	
XX	PR	19-DEC-2001; 2001US-0340873P.	
XX	PR	21-FEB-2002; 2002US-0357842P.	
XX	PR	21-FEB-2002; 2002US-0357843P.	
XX	PR	21-FEB-2002; 2002US-0357844P.	
XX	PR	15-MAR-2002; 2002US-0364134P.	
XX	PR	08-APR-2002; 2002US-0370144P.	
XX	PR	08-APR-2002; 2002US-0370206P.	
XX	PR	08-APR-2002; 2002US-0370247P.	
XX	PR	17-APR-2002; 2002US-0372794P.	
XX	PR	21-APR-2002; 2002US-0371679P.	
XX	XX	(GENE-) GENE LOGIC INC.	
XX	PA	Mendrick D, Porter M, Johnson K, Higgins B, Castle A, Elashoff M;	
XX	PI	WPI; 2003-148464/14.	
XX	DR	Predicting at least one toxic effect of a compound, useful for toxicity	
XX	XX	cell sample exposed to the compound, and comparing the gene expression	
XX	PT	profile to a database.	
XX	PT	Example 4; Page; 446pp; English.	
XX	PS	The invention relates to a novel method of predicting at least one toxic	
XX	XX	effect of a compound. The method comprises a gene expression profile of a	
XX	CC	tissue or cell sample exposed to the compound, and comparing the gene	
XX	CC	expression profile to a database comprising at least part of the data or	
XX	CC	information given in the specification. The methods are useful for	

CC	predicting at least one toxic effect of a compound, predicting the															
CC	progression of a toxic effect of a compound, predicting the renal															
CC	toxicity of a compound, or identifying toxicity markers in tissues or															
CC	cells exposed to known renal toxin. The genes are useful as toxicity															
CC	markers in drug screening and toxicity assays, in monitoring disease or															
CC	physiological states, or disease progression. This polynucleotide															
CC	represents a rat DNA sequence relating to the toxic effect database															
CC	described in the specification. NOTE: The sequence data for this patent															
CC	did not form part of the printed specification, but was obtained in															
CC	electronic format directly from the World Intellectual Property															
CC	Organization															
XX																
SQ	Sequence 4529 BP; 1018 A; 1179 C; 1187 G; 1145 T; 0 U; 0 Other;															
	Query Match	42.9%	Score	517.2;	DB	10;	Length	4529;								
	Best Local Similarity	78.5%;	Pred. No.	2.3e-143;												
	Matches	832;	Conservative	1;	Mismatches	201;	Indels	26;	Gaps	17;						
Qy	1	GC	GGCGT	GAGT	CGG	CCCC	CCCC	CGAGT	CACG	TGACG	GTGAC	TGCGGGGGT	TCTCCAC	TATC	60	
Db	3277	GC	CGCGGGG	AGCGGCC	CTCC	CGATG	ACG	TGACG	CGCGC	CACT	CGAGCGCT	C	CTGCTG	TGA	3335	
Qy	61	GC	TTAC	TACT	CTCC	CTCTG	CAG	GAAC	CCG	CGGAT	ATGG	CTGCGCTG	CCCCCG	CGCG	120	
Db	3336	GC	TG	CCCAT	CTCC	CGCTCC	CA	CAAT	CCG	CGGTG	ATGG	CTGCGCGCTT	CCTG	CGCGCGT	3395	
Qy	121	AT	TTCT	CTCC	CGCTG	CT	TCC	TTCTCT	CTCTGGG	CTT	CC	TCTCTCT	CTCCGCT	CCGATGG	179	
Db	3396	TT	CGCT	CTCC	CACTG	CTG	TC	CGTTC	CTCTTGG	CGCTG	CTG	CTCTCT	CTGCT	CATGG	3455	
Qy	180	CG	CAG	CGGCT	GC	CAC	CAAA	-	GGC	GCCTT	CC	CTGGAT	PAC	GGTCA	CTTTT	238
Db	3456	CG	CAG	CGGCT	GC	CAC	CAAGG	CGCG	CCCTT	CC	CTTGG	ACAC	AGTCA	CTTCT	TACAAGGT	3515
Qy	239	CAT	TCC	CA	AA	CAAG	TTC	GT	C	TG	TG	AGTT	TC	GC	AC	297
Db	3516	CAT	TCC	CA	AA	CAAG	TTC	GT	C	TG	TG	AGTT	TC	GC	AC	3575
Qy	298	GC	AGG	TGAG	TCT	CA	AGG	GTCT	-	T	CT	GAA	AAC	T	CG	356
Db	3576	GC	AAG	TGAG	TTC	CA	AGG	GTCT	CG	CTG	GAA	ACT	CAG	CTCC	AG	3635
Qy	357	AG	AGT	GGG	ATCT	C	A	GAT	TAT	-	G	TG	CA	AG	CTG	415
Db	3636	AG	AGT	GGG	ATCT	C	A	GAT	TAT	-	G	TG	CA	AG	CTG	3695
Qy	416	GCT	GG	CA	AA	GAG	AG	CT	T	AC	CA	-	T	CT	T	474
Db	3696	GCT	GG	CA	AA	GAG	AG	CT	T	AC	CA	-	T	CT	T	3755
Qy	475	AG	T	CC	CA	T	AC	T	AG	GG	CA	GT	-	AG	G	533
Db	3756	TG	T	CC	CA	T	AC	T	AG	GG	CA	GT	-	AG	G	3815
Qy	534	G	G	T	CT	ACT	AG	G	T	AT	G	CT	G	G	-	592
Db	3816	AG	T	CT	ACT	AG	G	T	AT	G	CT	G	G	-	3875	
Qy	593	C	A	G	G	CC	T	T	G	T	G	G	A	G	-	651
Db	3876	C	A	G	G	CC	T	T	G	T	G	G	A	G	-	3935
Qy	652	A	A	G	T	G	A	A	G	G	A	G	A	G	-	710
Db	3936	A	A	G	T	G	A	A	G	A	G	A	G	-	3995	
Qy	711	C	T	T	A	G	A	C	A	A	G	G	G	A	G	769
Db	3996	C	T	T	A	G	A	C	A	A	G	G	A	G	-	4055
Qy	770	G	A	A	G	A	C	A	A	G	A	G	G	A	G	829
Db	4056	G	A	A	G	A	C	A	A	G	A	G	G	A	G	4113

Qy	830	GCCTTCAGAAAGAGGGGGCCGAGAAAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCC	889
Db	4114	GCCTTCGCAAGAAAGCGCGCGAGAGGAGGAGCTGT-----GAGGGCACCAGGCTGCT	4167
Qy	890	AGGTTTGGTGGGGGTAGGAGGAGGGANAGTTAAACCTGCTGGCTGTGTGANTCCCTTGTGGAA	949
Db	4168	AGGTTTGTCTGAGGGCGGGAGAGG--AGACCCGCTGCTGACTGTGTGGGGCCCTGTGGG--	4223
Qy	950	TATAAGGGGYSKGGGAAAAGGGTACTAACCACGATTTCTGAGCCCTGAGTATGCCTG	1009
Db	4224	TGAGGGGGCAGTGGAGCAAGCAGTCTCTGAGCCAGAGACCTGTGCCCCGAG--TGCCTG	4281
Qy	1010	GACATTGATCTAACATGACCATGCTTGGGATGCTCTAG	1049
Db	4282	GACACTGAGGCTCTGAGAGCATACCTGGACATCCTTGG	4321
RESULT 11			
ADP72860			
ID	ADP72860 standard; DNA; 1139 BP.		
XX	AC	AC	ADP72860;
XX	XX	XX	26-AUG-2004 (first entry)
DT	Renal toxin progression gene marker #1449.		
XX	ds; toxic effect; gene expression profile; kidney tissue;		
KW	differential gene expression; toxicity progression; toxicity marker;		
KW	drug screening; toxicity assay; kidney pathology; nephritis;		
KW	kidney necrosis; glomerular injury; tubular injury;		
KW	focal segmental glomerulosclerosis.		
XX	XX		
OS	Rattus norvegicus.		
XX	XX		
PN	WO2004048598-A2.		
XX	XX		
PD	10-JUN-2004.		
XX	XX		
XX	24-NOV-2003; 2003WO-US037556.		
PF	XX		
XX	22-NOV-2002; 2002US-00301856.		
PR	XX		
XX	(GENE-) GENE LOGIC INC.		
PA	XX		
XX	XX		
PI	Mendrick DL, Porter MW, Johnson KR, Castle A, Higgs B;		
PI	Elashoff M;		
XX	XX		
XX	WPI; 2004-460771/43.		
DR	XX		
XX	Predicting (the progression of) a toxic effect of a compound, for		
PT	monitoring the progression of renal disease states, comprises preparing a		
PT	gene expression profile of a kidney tissue or cell sample exposed to the		
PT	compound.		
XX	XX		
PS	Claim 11; SEQ ID NO 1449; 265pp; English.		

The invention relates to a method of predicting (the progression of) a toxic effect of a compound by preparing a gene expression profile of a kidney tissue or cell sample exposed to the compound and comparing the gene expression profile to a database, or detecting the level of gene(s) expression in a tissue or cell sample exposed to the compound, where differential gene expression compared to a control indicates a toxic effect (toxicity progression). The method is useful for predicting (the progression of) at least one toxic effect of a compound. The genes are useful as toxicity markers in drug screening and toxicity assays. The methods are useful for predicting the likelihood that a compound or test agent will induce various specific kidney pathologies, such as nephritis, kidney necrosis, glomerular and tubular injury, or focal segmental glomerulosclerosis. The methods are useful for determining the similarity of a toxic response to one or more individual compounds and for predicting or elucidating the potential cellular pathways influenced.

QY 811 GAGCTTAACATCTGACCTGCTTCCAGAGAAGGGGCGGAGAAAGAGAGCTGTAAAA 870
DB 61 GAGCTTAACATCTGACCTGCTTCCAGAGAAGGGGCGGAGAAAGAGAGCTGTAAAA 120
QY 871 AGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGGAGTAACTCTGCTGG 930
DB 121 AGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGGGAGTAACTCTGCTGG 180
QY 931 CTGTGANTCCCTTGTGGATATATAAGGGGYSKGGGAAAGHGGTACTAACCCAGATTC 990
DB 181 CTGTGAGTCCCTTGTGGATATATAAGGGGTAGTGGGAAAGTGGTACTAACCCAGATTC 240
QY 991 TGAGCCCTGAGTATGCTGGACATTTGCTAACATGACCATGCTTGGGATGCTCTAGC 1050
DB 241 TGAGCCCTGAGTATGCTGGACATTTGCTAACATGACCATGCTTGGGATGCTCTAGC 300
QY 1051 TGGTCTGGGATAGCTGGAGCACTTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGA 1110
DB 301 TGGTCTGGGATAGCTGGAGCACTTTACTCAGTGGCTGGTGAATGACACCTCAGAAGGA 360
QY 1111 ATGAGTGTATAGAGAGGA-GAGAGGAGTACTGCCAGGCTTTTGACAGATGTAATTC 1169
DB 361 ATNAGTGTATAGAGAGGAGGAGGAGTACTGCCAGGCTTTTGACAGATGTAATTC 420
QY 1170 TCATTCAATTAATGAA-GTTTCAGTGTGTTTGGTTAA 1201
DB 421 TCATTCAATTAATGAGTTTCAGTGTGTTTGGTTAA 453

RESULT 13

AAC03762
ID AAC03762 standard; cDNA; 500 BP.

XX AC AAC03762;

XX DT 06-OCT-2000 (first entry)

XX DE Human secreted protein 5' EST, SEQ ID NO: 3760.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX FN EP1033401-A2.

XX PD 06-SEP-2000.

XX PF 21-FEB-2000; 2000EP-00200610.

XX PR 26-FEB-1999; 99US-0122487P.

XX PA (GBST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX DR P-ESDB; AAG03756.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX PT diagnostic, forensic, gene therapy and chromosome mapping procedures.

XX FS Claim 1; SEQ ID NO 3760; 71pp + Sequence Listing; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from
XX CC mRNAs encoding secreted proteins. An ORF has been identified within the
XX CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
XX CC derived from 30 different tissues. EST sequences usually correspond
XX CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
XX CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
XX CC well suited for isolating cDNA sequences derived from the 5' ends of
XX CC mRNAs and even in those cases where longer cDNA sequences have been

CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.

SQ Sequence 500 BP; 100 A; 155 C; 133 G; 112 T; 0 U; 0 Other;

Query Match 33.4%; Score 402.2; DB 3; Length 500;
Best Local Similarity 96.8%; Pred. No. 1.9e-109;
Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

QY 26 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACTCTCTCTCGAGAA 85
DB 1 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACTCTCTCTCGAGAA 60
QY 86 CCGGCGCATATAGGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCTT 144
DB 61 CCGGCGCATATAGGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTT 120
QY 145 CTCTGGGCTTCT 203
DB 121 CTCTGGGCTTCT 180
QY 204 GCCCTTCCCTTGGATACGCTCATTCTTCAAGGTCATTCCTCCAAAGCAAGTTCGTC-TG 262
DB 181 GCCCTTCCCTTGGATACGCTCATTCTTCAAGGTCATTCCTCCAAAGCAAGTTCGTC 240
QY 263 GTGAAGTTCGACACCCAGTACCTTACGTCGAGAGCAGGATGAGTTCAGAGGCTCT-CT 321
DB 241 GTGAAGTTCGACACCCAGTACCTTACGTCGAGAGCAGGATGAGTTCAGAGGCTCTTGT 300
QY 322 GAAAGTCTGGCTTCCAGGATGATCTCTTGTGGCAGAGGAGTGGGATCTCAGATTATGTG 381
DB 301 GAAAGTCTGGCTTCCAGGATGATCTCTTGTGGCAGAGGAGTGGGATCTCAGATTATGTG 360
QY 382 ACAAGCTGAACATGGAGCTGAGTGAAGAAATCAAGCTGGACAAAGAGAGCTACCCA-TCT 440
DB 361 GACAAGCGAACATGGAGCTGAGTGAAGAAATCAAGCTGGACAAAGAGAGCTACCCAAGTCT 420
QY 441 TCTACCTCTTCCGGGATGGGACATTTGAGACCCAGTCCCATACACTGGGCGAGTT--AG 498
DB 421 TCTACCTCTTCCGGGATGGGACATTTGAGAA-CCAGTCCCATACACTGGGCGAGTTAAG 479
QY 499 GTTGGAGCCATCCAGCGCTGG 519
DB 480 GTTGGAGCCATCCAGCGCTGG 500

RESULT 14

AAS83797
ID AAS83797 standard; cDNA; 2087 BP.

XX AC AAS83797;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #19601.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX FN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX XX 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

PI Jiang Y, Chenault RA, Xu J, Indirias CY, Lodes MJ, Secrist H;
PI Carter D, Fanger GR, Smith CL, Durham M, Stolk JA;
XX WPI; 2003-040540/03.
DR
XX
XX New isolated nucleic acids and polypeptides capable of eliciting a
PT humoral and/or cellular immune response, useful for diagnosing,
PT preventing or treating cancer, particularly colon cancer.
XX
PS Claim 1; SEQ ID NO 1394; 87pp; English.
XX
XX The invention relates to polynucleotide and polypeptide sequences
CC associated with cancer, particularly colon cancer. Also disclosed are (i)
CC an expression vector comprising the polynucleotide, (ii) a host cell
CC transformed or transfected with the expression vector, (iii) an isolated
CC antibody, or its antigen-binding fragment, which specifically binds to
CC the polypeptide, (iv) a method of detecting or determining the presence
CC of cancer in a patient, (v) a fusion protein comprising at least one of
CC the polypeptides, (vi) an oligonucleotide that hybridizes to the
CC polynucleotide sequence under highly stringent conditions, and (vii) a
CC method of stimulating and/or expanding T cells specific for a tumour
CC protein. The polypeptide specifically comprises the amino acid sequence
CC of C634S, C635S, C637S, C640S, C636S or one of the potential open reading
CC frames (ORFs) of C636S. These polypeptides are encoded by the
CC polynucleotide sequences, where both are capable of eliciting a humoral
CC and/or cellular immune response. The polynucleotides, polypeptides, and
CC antibodies are useful for diagnosing, preventing or treating cancer,
CC particularly colon cancer. The polynucleotide and polypeptide sequences
CC are also useful in DNA strand invasion, antisense inhibition, mutational
CC analysis, nucleic acid purification, isolation of transcriptionally
CC active genes, blocking or transcription factor binding, genome cleavage
CC or in situ hybridisation, and as enhancers of transcription or
CC biomarkers. This sequence represents a human colon cancer associated
CC cDNA. Note: The sequence data for this patent was obtained in electronic
CC format directly from the USPTO web site at seqdata.uspto.gov
XX
SQ Sequence 395 BP; 106 A; 88 C; 128 G; 70 T; 0 U; 3 Other;

Query Match 27.3%; Score 328.8; DB 11; Length 395;
Best Local Similarity 97.5%; Pred. No. 1.5e-87;
Matches 385; Conservative 0; Mismatches 5; Indels 5; Gaps 5;

QY 506 CCATCCAGCGCTGGCTGAAGGGGCAAGGGCTACCTAGGTATGCTGG-TGCCCTGCCTG 564
Db 1 CCATCCAGCGCTGGCTGAAGGGGCAAGGGCTACCTAGGTATGCTGGCTGCCTG 60
QY 565 TATACGAGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGGTGTGAGG-CGCGCCAGGCC 623
Db 61 TATACGAGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGGTGTGAGGCCCGCAGGCC 120
QY 624 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAG-AGTGGGCG 682
Db 121 TCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCANAAGAAGTGGGCG 180
QY 683 AGCATACCTGAGATCATGGGAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAG 741
Db 181 AGCAATACCTGAGATCATGGGAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAG 240
QY 742 AGATGACACGGATCCCGAGGCTGATTGAGAAGAACAGATGATGAC-GGCAGAGAGAGG 800
Db 241 AGATGACACGGATCCCGAGGCTGATTGAGAAGAACAGATGATGATCGGAGAGAGAGG 300
QY 801 AGCTCCAGAGAGCTTAAACATCTCTGACTGCTCCAGAAGAGGGGCCGAGAAAGAGG 860
Db 301 AGCTCCAGAGAGCTTAAACATCTCTGACTGCTCCAGAAGAGGGGCCGAGAAAGAGG 360
QY 861 AGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTT 895
Db 361 AGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTT 395

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 190.034 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-7
Perfect score: 1205
Sequence: 1 GCGGCGTGAGTCGCCGCC.....TCAGTGTGTTGGTTAAGTGG 1205

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:**

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:**
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:**
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:**
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:**
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:**
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1201	99.7	1205	2	Sequence 7, Appli
2	1201	99.7	1205	4	Sequence 7, Appli
3	810.8	67.3	962	4	Sequence 4780, Ap
4	742.8	61.6	1364	4	Sequence 16522, A
5	402.2	33.4	500	4	Sequence 3760, Ap
6	400.4	33.2	601	4	Sequence 169883,
7	191.2	15.9	363	4	Sequence 8299, Ap
8	70.8	5.9	7218	1	Sequence 14, Appl
9	52	4.3	7218	1	Sequence 14, Appl
10	42.2	3.5	43267	4	Sequence 17117, A
11	37.8	3.1	289	3	Sequence 17, Appl
12	37.8	3.1	289	3	Sequence 17, Appl
13	37.8	3.1	289	3	Sequence 17, Appl
14	37.2	3.1	1526	3	Sequence 14464, A
15	37.2	3.1	1526	3	Sequence 63, Appl
16	37	3.1	27630	4	Sequence 17314, A
17	37	3.1	27684	4	Sequence 16362, A
18	36.8	3.1	53526	3	Sequence 12504, A
19	36.8	3.1	53577	3	Sequence 2, Appli
20	36.6	3.0	4749	4	Sequence 189, App
21	36.4	3.0	36954	4	Sequence 15304, A
22	36.2	3.0	35688	4	Sequence 16873, A
23	36.2	3.0	68580	4	Sequence 15844, A
24	35.6	3.0	187595	4	Sequence 15446, A
25	35.4	2.9	2810	3	Sequence 6, Appli
26	35.4	2.9	3225	1	Sequence 45, Appl
27	35.4	2.9	3225	5	Sequence 91, Appl

C 28 35.2 2.9 296 4 US-09-216-393B-97 Sequence 97, Appl
C 29 35.2 2.9 20662 4 US-09-949-016-16717 Sequence 16717, A
C 30 35.2 2.9 43435 4 US-09-949-016-12909 Sequence 12909, A
31 35 2.9 364 4 US-09-621-376-17202 Sequence 17202, A
32 35 2.9 2338 4 US-09-582-337-1 Sequence 1, Appli
33 35 2.9 767677 4 US-09-949-016-12147 Sequence 12147, A
34 35 2.9 767677 4 US-09-949-016-17361 Sequence 17361, A
35 34.8 2.9 375 4 US-09-513-999C-13614 Sequence 13614, A
36 34.8 2.9 9168 4 US-09-687-731-11 Sequence 11, Appl
37 34.4 2.9 289 3 US-09-007-005-17 Sequence 17, Appl
38 34.4 2.9 289 3 US-09-244-796-17 Sequence 17, Appl
39 34.4 2.9 601 4 US-09-949-016-51960 Sequence 51960, A
40 34.4 2.9 12368 4 US-09-949-016-13243 Sequence 13243, A
41 34.4 2.9 247299 4 US-09-949-016-17590 Sequence 17590, A
C 42 34.2 2.8 430 4 US-09-621-976-16656 Sequence 16656, A
43 34.2 2.8 62386 4 US-09-949-016-12823 Sequence 12823, A
44 34.2 2.8 390890 4 US-09-949-016-14720 Sequence 14720, A
C 45 34 2.8 39085 4 US-09-949-016-14479 Sequence 14479, A

ALIGNMENTS

RESULT 1
US-08-680-395-7
; Sequence 7, Application US/08680395
; Patent No. 5892010
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W.
; APPLICANT: Collins, Colin
; APPLICANT: Hwang, Soo-in
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowbel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: Genes from the 20q13 Amplicon and Their
; TITLE OF INVENTION: Uses
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/680,395
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 023070-06890005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: 1..1205
; LOCATION: /note= "cdna clone 1b4 for a serine
; OTHER INFORMATION: threonine kinase"
; OTHER INFORMATION:
US-08-680-395-7

Query Match		99.7%;	Score 1201;	DB 2;	Length 1205;				
Best Local Similarity		100.0%;	Pred. No. 0;						
Matches 1205; Conservative		0;	Mismatches	0;	Indels	0;	Gaps	0;	
Qy	1	CGCGCGTGAGTCCGCCCCCAGCTACGTCACCGCTGACCTCGGGGGGTTCTCCACATATC	60						
Db	1	CGCGCGTGAGTCCGCCCCCAGCTACGTCACCGCTGACCTCGGGGGGTTCTCCACATATC	60						
Qy	61	GCTTACCTACCTCCCTCTGCAGGAACCGCGGATATGCTGCCCTGTCGCCCGCGCGC	120						
Db	61	GCTTACCTACCTCCCTCTGCAGGAACCGCGGATATGCTGCCCTGTCGCCCGCGCGC	120						
Qy	121	ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTCTCGCTCCCTCGCATGCG	180						
Db	121	ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTCTCGCTCCCTCGCATGCG	180						
Qy	181	GGCAGCGGCTGCAACCAAGCGGCTTCCCTGGATACGGTCACTTCTTACAAGGTCA	240						
Db	181	GGCAGCGGCTGCAACCAAGCGGCTTCCCTGGATACGGTCACTTCTTACAAGGTCA	240						
Qy	241	TTCCCAAAGCAAGTTCGTCTGAGTTCGAGTTCGACACCCAGTACCCTACGGTGAGAGCA	300						
Db	241	TTCCCAAAGCAAGTTCGTCTGAGTTCGAGTTCGACACCCAGTACCCTACGGTGAGAGCA	300						
Qy	301	GGATGAGTTCAAGCGTCTTCTGAAACCTCGGCTTCCAGCGATGATCTCTTGGTGCGAGAG	360						
Db	301	GGATGAGTTCAAGCGTCTTCTGAAACCTCGGCTTCCAGCGATGATCTCTTGGTGCGAGAG	360						
Qy	361	GTGGGATCTCAGATTATGTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGG	420						
Db	361	GTGGGATCTCAGATTATGTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGG	420						
Qy	421	ACAAAGAGAGTACCCATCTTCTACCTCTCCGGATGGGACTTTGAGACCCAGTCCC	480						
Db	421	ACAAAGAGAGTACCCATCTTCTACCTCTCCGGATGGGACTTTGAGACCCAGTCCC	480						
Qy	481	ATACACTGGGCGAGTGGGAGCCATCCAGCGCTGGCTGAAAGGGGCAAGGGTCTAC	540						
Db	481	ATACACTGGGCGAGTGGGAGCCATCCAGCGCTGGCTGAAAGGGGCAAGGGTCTAC	540						
Qy	541	CTAGTATGCTGTGCTGCTGTATACGACGCTTGGCCGGGGAGTTATACGGGCT	600						
Db	541	CTAGTATGCTGTGCTGCTGTATACGACGCTTGGCCGGGGAGTTATACGGGCT	600						
Qy	601	CTGTGTGGAGCGCCAGGCGCTTCTGAAGCAGGGCAAGTAACTCTCAAGTGTGAA	660						
Db	601	CTGTGTGGAGCGCCAGGCGCTTCTGAAGCAGGGCAAGTAACTCTCAAGTGTGAA	660						
Qy	661	GGAGACTCAGAAAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAA	720						
Db	661	GGAGACTCAGAAAGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAA	720						
Qy	721	GGGAGACATTTCCAGCATCAGAGATGACACCGATCGCCAGGCTGATGAGAAACAGA	780						
Db	721	GGGAGACATTTCCAGCATCAGAGATGACACCGATCGCCAGGCTGATGAGAAACAGA	780						
Qy	781	TGAGTGCGGAGAGGAGGCTCCAGAGCTTAAACATCTGAGTCCCTGCTCCAGAA	840						
Db	781	TGAGTGCGGAGAGGAGGCTCCAGAGCTTAAACATCTGAGTCCCTGCTCCAGAA	840						
Qy	841	GAAGGGGCGCAGAAAGAGGAGCTGTAAAGGCTGTCTGTATTTCCAGGGTTTGGTG	900						
Db	841	GAAGGGGCGCAGAAAGAGGAGCTGTAAAGGCTGTCTGTATTTCCAGGGTTTGGTG	900						
Qy	901	GGGTAGGGAAGGAGATTAACCTGTGGCTGTGANTCCCTTGTGGAATATAAGGGGY	960						
Db	901	GGGTAGGGAAGGAGATTAACCTGTGGCTGTGANTCCCTTGTGGAATATAAGGGGY	960						
Qy	961	MSKGGGAAAGWGTACTAAACCAAGATTTGAGCCCTGAGTATGCCCTGGACATTGATGC	1020						
Db	961	MSKGGGAAAGWGTACTAAACCAAGATTTGAGCCCTGAGTATGCCCTGGACATTGATGC	1020						

Qy	1021	TAAATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGAGACATTACTCA	1080						
Db	1021	TAAATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGAGACATTACTCA	1080						
Qy	1081	GCTGGCTGTGAAATGACACCTCAGAGGAATGAGTCTATAGAGGAGAGAGGAGTGT	1140						
Db	1081	GCTGGCTGTGAAATGACACCTCAGAGGAATGAGTCTATAGAGGAGAGAGGAGTGT	1140						
Qy	1141	ACTGCCCAGCTCTTTGACAGATGTAATCTCATTTCAATTAAGTTTCAGTGTTTGGTTA	1200						
Db	1141	ACTGCCCAGCTCTTTGACAGATGTAATCTCATTTCAATTAAGTTTCAGTGTTTGGTTA	1200						
Qy	1201	AGTGG 1205							
Db	1201	AGTGG 1205							
RESULT 2									
US-08-892-695-7									
; Sequence 7, Application US/08892695A									
; Patent No. 6808878									
; GENERAL INFORMATION:									
; APPLICANT: Gray, Joe W									
; APPLICANT: Collins, Collin									
; APPLICANT: Hwang, Soo In									
; APPLICANT: Godfrey, Tony									
; APPLICANT: Kowel, David									
; APPLICANT: Rommens, Johanna									
; TITLE OF INVENTION: GENES FROM THE 20013 AMPLICON AND THEIR USES									
; FILE REFERENCE: 2500.124US3									
; CURRENT APPLICATION NUMBER: US/08/892,695A									
; CURRENT FILING DATE: 1997-07-15									
; EARLIER APPLICATION NUMBER: 08/785,532									
; EARLIER FILING DATE: 1997-01-17									
; EARLIER APPLICATION NUMBER: 08/731,499									
; EARLIER FILING DATE: 1996-10-16									
; EARLIER APPLICATION NUMBER: 08/680,395									
; EARLIER FILING DATE: 1996-07-15									
; NUMBER OF SEQ ID NOS: 59									
; SOFTWARE: PatentIn Ver. 2.0									
; SEQ ID NO 7									
; LENGTH: 1205									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence:1b4									
; FEATURE:									
; NAME/KEY: modified_base									
; LOCATION: (916)									
; OTHER INFORMATION: n is A, C, G, T, or U									
; FEATURE:									
; NAME/KEY: modified_base									
; LOCATION: (937)									
; OTHER INFORMATION: n is A, C, G, T, or U									
; US-08-892-695-7									
Query Match									
Best Local Similarity 99.7%; Score 1201; DB 4; Length 1205;									
Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CGCGCGTGAGTCCGCCCCCAGCTACGTCACCGCTGACCTCGGGGGTTCACATATC	60						
Db	1	CGCGCGTGAGTCCGCCCCCAGCTACGTCACCGCTGACCTCGGGGGTTCACATATC	60						
Qy	61	GCTTACCTACCTCCCTCTGCAGGAACCGGGGATATGGCTGCCCTGCGCGCGCGC	120						
Db	61	GCTTACCTACCTCCCTCTGCAGGAACCGGGGATATGGCTGCCCTGCGCGCGCGC	120						
Qy	121	ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTCTCGCTCCCTCGCATGCG	180						
Db	121	ATTTCTCTCCCGCTGCTTCCCTTCTCCTGGGCTTCTCTCGCTCCCTCGCATGCG	180						
Qy	181	GGCAGCGGCTGCAACCAAGCGGCTTCCCTGGATACGGTCACTTCTTACAAGGTCA	240						
Db	181	GGCAGCGGCTGCAACCAAGCGGCTTCCCTGGATACGGTCACTTCTTACAAGGTCA	240						

```
Db 181 GCACGGGCTGCGACACCAAGCGCCCTTCCCTGGATACGGTCACTTCTACAGGTCA 240
QY 241 TTCCCAAAAGCAAGTTGCTGCTGGTGAAGTTGACACCCAGTACCCCTACGGTGAAGA 300
Db 241 TTCCCAAAAGCAAGTTGCTGCTGGTGAAGTTGACACCCAGTACCCCTACGGTGAAGA 300
QY 301 GGATGAGTTCAAGCGCTCTTGAAACCTCGGCTCCAGCGATGATCTCTTGGTGCAGAG 360
Db 301 GGATGAGTTCAAGCGCTCTTGAAACCTCGGCTCCAGCGATGATCTCTTGGTGCAGAG 360
QY 361 GTGGGGATCTCAGATTATGTGAACAGCTGAACATGGAGCTGAGTGAGAAATACAGAGCTGG 420
Db 361 GTGGGGATCTCAGATTATGTGAACAGCTGAACATGGAGCTGAGTGAGAAATACAGAGCTGG 420
QY 421 ACAAGAGAGCTACCAATCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCC 480
Db 421 ACAAGAGAGCTACCAATCTTCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCC 480
QY 481 ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGTCTAC 540
Db 481 ATACACTGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGTCTAC 540
QY 541 CTAGTATGCTGCTGCTGCTGCTATACGACGCTGCGCGGGAGTTCTACAGGGCT 600
Db 541 CTAGTATGCTGCTGCTGCTGCTATACGACGCTGCGCGGGAGTTCTACAGGGCT 600
QY 601 CTGTTGTGGAGCCGCCAGGCGCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAA 660
Db 601 CTGTTGTGGAGCCGCCAGGCGCTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAA 660
QY 661 CGAGACTCAGAAGAGTGGGCGAGCAATACCTGAAGATCA TGGGGAAGATCTTAGACCA 720
Db 661 CGAGACTCAGAAGAGTGGGCGAGCAATACCTGAAGATCA TGGGGAAGATCTTAGACCA 720
QY 721 GGGGAGCACTTCCAGCATCAGAGATGACACGATCGCAGGCTGATCAGAGAAACAAGA 780
Db 721 GGGGAGCACTTCCAGCATCAGAGATGACACGATCGCAGGCTGATCAGAGAAACAAGA 780
QY 781 TGAGTGACGGCAGAAAGGAGGAGCTCCAGAAGAGCTTAAACATCTCTGACTGCTCCAGAA 840
Db 781 TGAGTGACGGCAGAAAGGAGGAGCTCCAGAAGAGCTTAAACATCTCTGACTGCTCCAGAA 840
QY 841 GAAGGGGCGGAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
Db 841 GAAGGGGCGGAGAAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTGGTG 900
QY 901 GGGGTAGGGAGGGAGGAGTTAACTGCTGGCTGCTGANTCCCTTGTGGAAATATAAGGGGY 960
Db 901 GGGGTAGGGAGGGAGGAGTTAACTGCTGGCTGCTGANTCCCTTGTGGAAATATAAGGGGY 960
QY 961 MSKGGGAAAAGWGGTACTAACCCAGATTTGAGCCCTGAGTATGCTGAGCATTTGATGC 1020
Db 961 MSKGGGAAAAGWGGTACTAACCCAGATTTGAGCCCTGAGTATGCTGAGCATTTGATGC 1020
QY 1021 TAAATGACCATGCTTGGGATGTCTAGCTGGTCTGGGGATGAGCTGAGAGCATTTACTCA 1080
Db 1021 TAAATGACCATGCTTGGGATGTCTAGCTGGTCTGGGGATGAGCTGAGAGCATTTACTCA 1080
QY 1081 GGTGGCTGGTGAATGACACCTCAGAGCAATGAGTGTATAGAGAGAGAGAGGTGT 1140
Db 1081 GGTGGCTGGTGAATGACACCTCAGAGCAATGAGTGTATAGAGAGAGAGAGGTGT 1140
QY 1141 ACTGCCCAAGGCTTTTGACAGATGTAAATCTCAATCAATTAAGTTTCAAGTGTGGTTA 1200
Db 1141 ACTGCCCAAGGCTTTTGACAGATGTAAATCTCAATCAATTAAGTTTCAAGTGTGGTTA 1200
QY 1201 AGTGG 1205
Db 1201 AGTGG 1205
```

RESULT 3

US-09-949-016-4780

; Sequence 4780, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-10-03

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4780

; LENGTH: 962

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4780

Query Match 67.3%; Score 810.8; DB 4; Length 962;

Best Local Similarity 97.8%; Pred. No. 1.9e-233;

Matches 941; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

```
QY 86 CCCGCGGATATGGCTGCGCTGTCGCCCGCGCGCGCATTTCTCCCCGCTGCTT-CCCTT 144
Db 1 CCCGCGGATATGGCTGCGCTGTCGCCCGCGCGCGCATTTCTCCCCGCTGCTTCCCTT 60
QY 145 CTCCTGGGCTTCTGCTCTCTCCGCTCCGATGGCGGACGGGCTGCACACCAA-GGC 203
Db 61 CTCCTGGGCTTCTGCTCTCTCCGCTCCGATGGCGGACGGGCTGCACACCAAGGCG 120
QY 204 GCCCTTCCCCTGGATACGGTCACTTCTACAGGTCATTTCCCAAGCAAGTTCGTC-TG 262
Db 121 GCCCTTCCCCTGGATACGGTCACTTCTACAGGTCATTTCCCAAGCAAGTTCGTC-TG 180
QY 263 GTGAAGTTCGACACCCAGTACCCCTACGGTGAGAGAGGATGAGTTCAGGCTCTT-CT 321
Db 181 GTGAAGTTCGACACCCAGTACCCCTACGGTGAGAGAGGATGAGTTCAGGCTCTTGT 240
QY 322 GAAACTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTAT-GT 380
Db 241 GAAACTCGGCTTCCAGCGATGATCTCTTGTGGCAGAGTGGGATCTCAGATTATGCT 300
QY 381 GACAGCTGAACATGGAGCTGAGTGAGAAATACAGCTGGGCAAGAGAGCTACCCA-TC 439
Db 301 GACAGCTGAACATGGAGCTGAGTGAGAAATACAGCTGGGCAAGAGAGCTACCCAAGTC 360
QY 440 TTCTACCTCTTCCGGGATGGGACTTTTGAGAAACCCAGTCCCATACACTGGGGCAGTT-AG 498
Db 361 TTCTACCTCTTCCGGGATGGGACTTTTGAGAAACCCAGTCCCATACACTGGGGCAGTTAAG 420
QY 499 GTTGGAGCAATCCAGCGCTGGCTCAAGGGGCAAGGGGTCTACCTAGGTATGCTGG-TGC 557
Db 421 GTTGGAGCAATCCAGCGCTGGCTCAAGGGGCAAGGGGTCTACCTAGGTATGCTGGTTGC 480
QY 558 CTGCTGTATACGACGCGCTGGCGGGGAGTTCTACAGGGCTCTGCTGTGGAGG-CGCG 616
Db 481 CTGCTGTATACGACGCGCTGGCGGGGAGTTCTACAGGGCTCTGCTGTGGAGGCGCGC 540
QY 617 CAGGCCCTTTGAAGAGGGGCAAGATAAAGTCTCAAGTGTGAAGGAGACTCAGAAG-AG 675
Db 541 CAGGCCCTTTGAAGAGGGGCAAGATAAAGTCTCAAGTGTGAAGGAGACTCAGAAGAG 600
QY 676 TGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCATT-CCA 734
Db 601 TGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCATTCCCA 660
QY 735 GCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAAGCAAGATGAGTGAC-GGCAG 793
```

Db 661 GCATCAGAGATGACACGGATCGCCAGGCTGATTTGAGAAAGAACAGATGAGTGCAGGGAAG 720
Qy 794 AAGGAGGAGCTCCAGAGAGCTTAAACATCCTGACTGCTTCCAGAAAGGGGCGGAG 853
Db 721 AAGGAGGAGCTCCAGAGAGCTTAAACATCCTGACTGCTTCCAGAGAGAGGGGCGGAG 780
Qy 854 AAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGG 913
Db 781 AAGAGGAGCTGTAAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGAGGG 840
Qy 914 GANAGTTAACCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYSKSGGAAAAAGWG 973
Db 841 GAGAGTTAACCTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGGTAGTGGGAAAAGTG 900
Qy 974 GTACTAACCCACGATTTCTGAGCCCTGAGTATGCTGGACATTTGATGCTTAACATGACCANTG 1033
Db 901 GTACTAACCCACGATTTCTGAGCCCTGAGTATGCTGGACATTTGATGCTTAACATGACCANTG 960
Qy 1034 CT 1035
Db 961 CT 962
RESULT 4
US-09-949-016-16522
; Sequence 16522, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16522
; LENGTH: 13364
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(13364)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16522
Query Match 61.6%; Score 742.8; DB 4; Length 13364;
Best Local Similarity 98.1%; Pred. No. 2.6e-212;
Matches 818; Conservative 5; Mismatches 4; Indels 7; Gaps 7;
Qy 379 GTGCAAGCTGAACATGAGCTGAGTGAAGATACAAAGCTGGACAAAGAGAGCTACCCA- 437
Db 10700 GTGCAAGCTGAACATGAGCTGAGTGAAGATACAAAGCTGGACAAAGAGAGCTACCCA- 10759
Qy 438 TCTTCTACCTCTTCGGGATGGGACTTTTGAAGCCCAAGTCCCATACACTGGGGCAGTT- 496
Db 10760 TCTTCTACCTCTTCGGGATGGGACTTTTGAAGCCCAAGTCCCATACACTGGGGCAGTTA 10819
Qy 497 AGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTGG-T 555
Db 10820 AGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGTATGCTGGTT 10879
Qy 556 GCCTCCCTGTATACGACGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGTTGTGGAGG-CC 614
Db 10880 GCCTCCCTGTATACGACGCCCTGGCCGGGAGTTTCATCAGGGCCCTCTGTTGTGGAGGCC 10939

Qy 615 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACTCTCAAGTGTGAAGAGACTCAGAAG- 673
Db 10940 GCCAGGCCCTCTTGAAGCAGGGGCAAGATAACTCTCAAGTGTGAAGAGACTCAGAAGA 10999
Qy 674 AGTGGGCGGAGCATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-C 732
Db 11000 AGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCC 11059
Qy 733 CAGCATCAGAGATGACACGGATCCAGAGGCTGATTTGAGAAGAAACAAGATGAGTGAC-GGC 791
Db 11060 CAGCATCAGAGATGACACGGATCCAGAGGCTGATTTGAGAAGAAACAAGATGAGTGACGGGA 11119
Qy 792 AGAAGGAGGAGCTCCAGAGAGCTTAAACATCTCGACTGCTCCAGAAAGGGGGCCG 851
Db 11120 AGAAGGAGGAGCTCCAGAGAGCTTAAACATCTCGACTGCTCCAGAAAGGGGGCCG 11179
Qy 852 AGAAGGAGGAGCTTAAAGAGGCTGCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAG 911
Db 11180 AGAAGGAGGAGCTTAAAGAGGCTGCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAG 11239
Qy 912 GGGANAGTTAACCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYSKSGGAAAAAG 971
Db 11240 GGGAGAGTTAACCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYSKSGGAAAAAG 11299
Qy 972 WGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTGGACATTTGATGCTTAACATGACCA 1031
Db 11300 TGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCTGGACATTTGATGCTTAACATGACCA 11359
Qy 1032 TGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTG 1091
Db 11360 TGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTG 11419
Qy 1092 AAATGACACCTTCAGAGGAATGAGTGTCTATAGAGAGAGAGAGGTGTACTGCCAGGT 1151
Db 11420 AAATGACACCTTCAGAGGAATGAGTGTCTATAGAGAGAGAGAGGTGTACTGCCAGGT 11479
Qy 1152 CTTTGACAGATGTAATTTCTCAATTAAGTTTCAAGTTTGGTTAAAGTG 1205
Db 11480 CTTTGACAGATGTAATTTCTCAATTAAGTTTCAAGTTTGGTTAAAGTG 11533

RESULT 5
US-09-513-999C-3760
; Sequence 3760, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 3760
; LENGTH: 500
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 70..378
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 70..129
; OTHER INFORMATION: score 6.9
US-09-513-999C-3760
Query Match 33.4%; Score 402.2; DB 4; Length 500;

Best Local Similarity 96.8%; Pred. No. 1.2e-110;
Matches 485; Conservative 0; Mismatches 8; Indels 8; Gaps 7;

QY 26 CACGTGACCGTGAATCGGGGGGTTCTCCACTATCGCTTACCTACTCTCCCTCTGCAGAA 85
DB 1 CACGTGACCGTGAATCGGGGGGTTCTCCACTATCGCTTACCTACTCTCCCTCTGCAGAA 60
QY 86 CCCGCGGATATGGCTGCGCTGTGCGCGCGCGCAATTTCTCTCCCGCTGCTT-CCCTT 144
DB 61 CCCGCGGATATGGCTGCGCTGTGCGCGCGCGCAATTTCTCTCCCGCTGCTTCCCTT 120
QY 145 CTCCTGGGCTTCTCTCTCTCTCTCCGCTCGCATGCGCGCGCGCTGCAACAA-GGC 203
DB 121 CTCCTGGGCTTCTCTCTCTCTCTCCGCTCGCATGCGCGCGCGCTGCAACAAAGGC 180
QY 204 GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCTATCCAAAGCAAGTTGTC-TG 262
DB 181 GCCCTTCCCTGGATACGGTCACTTTCTAAGGTCTATCCAAAGCAAGTTGTCCTTG 240
QY 263 GTGAAGTTTCGACACCCAGTACCCCTACCGTGAAGCGAGGATGATTCAAGGCTCTT-CT 321
DB 241 GTGAAGTTTCGACACCCAGTACCCCTACCGTGAAGCGAGGATGATTCAAGGCTCTTCT 300
QY 322 GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 381
DB 301 GAAACTCGGCTTCCAGCGATGATCTTTGGTGGCAGAGTGGGATCTCAGATTATGTG 360
QY 382 ACAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCA-TCT 440
DB 361 GACAGCGAACAATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGCTACCCAGTCT 420
QY 441 TCTACTCTTCCGGATGGGAGCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT--AG 498
DB 421 TCTACTCTTCCGGATGGGAGCTTTGAGAA-CCAGTCCCATACACTGGGCGAGTTAAAG 479
QY 499 GTTGAGCCATCCAGCGCTGG 519
DB 480 GTTGAGCCATCCAGCGCTGG 500

RESULT 6
US-09-949-016-169883
; Sequence 169883, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169883
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-169883

Query Match 33.2%; Score 400.4; DB 4; Length 601;
Best Local Similarity 98.0%; Pred. No. 4.6e-110;
Matches 478; Conservative 1; Mismatches 12; Indels 7; Gaps 7;

QY 379 GTGACAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGCAAAAGAGAGCTACCCA- 437
DB 114 GTGACAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGCAAAAGAGAGCTACCCA 173

QY 438 TCTTCTACCTCTTCGGGATGGGACCTTTGAGAACCCAGTCCCATACACTGGGCGAGTT- 496
DB 174 TCTTCTACCTCTTCGGGATGGGACCTTTGAGAACCCAGTCCCATACACTGGGCGAGTTA 233
QY 497 AGGTTGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGGTCTACCTAGGTATGCTGG-T 555
DB 234 AGGTTGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGGTCTACCTAGGTATGCTGGTT 293
QY 556 GCCTGCTGTATACGACGCGCTGGCGGGGAGTTCATCAGGGCTCTCTGGTGTGAGG-CC 614
DB 294 GCCTGCTGTATACGACGCGCTGGCGGGGAGTTCATCAGGGCTCTCTGGTGTGAGGCGCC 353
QY 615 GCCAGGCCCTCTCAAGCAGGCGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG- 673
DB 354 GCCAGGCCCTCTCAAGCAGGCGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAGA 413
QY 674 AGTGGGCGGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-C 732
DB 414 AGTGGGCGGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCC 473
QY 733 CAGCATCAGATGACACGCGATCCGCGGCTGATTTGAGAGAAACAAGATGATGAC-GGC 791
DB 474 CAGCATCAGATGACACGCGATCCGCGGCTGATTTGAGAGAAACAAGATGATGACGCGGA 533
QY 792 AGAAGGAGGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGG 851
DB 534 AGAAGGAGGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGG 593
QY 852 AGAAGAG 859
DB 594 AGAAGAG 601

RESULT 7
US-09-513-999C-8299
; Sequence 8299, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513.999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36881
; SOFTWARE: Patent.pm
; SEQ ID NO 8299
; LENGTH: 363
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 175
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 176
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 188
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 214
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 215

; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-8299

Query Match 15.9%; Score 191.2; DB 4; Length 363;
Best Local Similarity 94.3%; Pred. No. 3.1e-47; Indels 2; Gaps 2;
Matches 215; Conservative 3; Mismatches 8;
Qy 26 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACCTACCTCCCTCTCGAGGAA 85
Db 1 CACGTGACCGCTGACTCGGGGGTTCCTCCACTATCGCTTACCTACCTCCCTCTCGAGGAA 60
Qy 86 CCGGGGATATGGCTGCGCGCTGTGCGCGCGCGCGCGGATTTCTCTCCCGCTGCTT-CCCTT 144
Db 61 CCGGGGATATGGCTGCGCGCTGTGCGCGCGCGCGGATTTCTCTCCCGCTGCTTCCCGCTT 120
Qy 145 CTCCTGGGCTTCCTGCTCTCTCGCTCGGATGCGCGCGGCTTCACACAA-GGC 203
Db 121 CTCCTGGGCTTCCTGCTCTCTCGCTCGGATGCGCGCGGCTTCACACAA-GGC 180
Qy 204 GCCCTTCCTCGGATACGGTCACTTTCTACAAGGTCACTTCCCAAGC 251
Db 181 GCCCTTCCTCGGATACGGTCACTTTCTACAAGGTATGGTGACAAGC 228

RESULT 8

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls

US-08-232-463-14

Query Match 5.9%; Score 70.8; DB 1; Length 7218;
Best Local Similarity 3.2%; Pred. No. 3e-10;
Matches 12; Conservative 228; Mismatches 132; Indels 0; Gaps 0;
Qy 628 GAACGAGGGCAAGATACCTCTCAAGTGTCAAGGAGACTCAGAAGAGTGGCCGAGCAA 687
Db 1417 RRR 1358
Qy 688 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCAGCTTCCAGCATCAGAGATGA 747
Db 1357 RRR 1298
Qy 748 CACGATCGCAGCGCTGATTGAGAAGAACAGATGATGACGCGAGAGAGAGCTCCA 807
Db 1297 RRR 1238
Qy 808 GAAGAGCTTAAACATCTGACTGCTTCCAGAAAGGGGCGGAGAGAGGAGCTGTA 867
Db 1237 RRR 1178
Qy 868 AAAAGGCTGCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGANAGTTAACCTGC 927
Db 1177 RRR 1118
Qy 928 TGGCTGTGANTCCCTGTGGAATATAAGGGGGMKGGGAAAGGWTACTAACCCAGCA 987
Db 1117 RRR 1058
Qy 988 TTCTGAGCCCTG 999
Db 1057 TCCTCGACCTG 1046

RESULT 9

US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:

```

;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

```

Query Match 4.3%; Score 52; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.00014;
Matches 10; Conservative 155; Mismatches 85; Indels

<i>Qy</i>	12	TCGCGCCCCCAGTACGCTGACTCGGGCGTTCTCCCACTATCGGTACTCTACC	71
<i>Db</i>	1200	Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y:::	1259
<i>Qy</i>	72	TCCTCTGCAGGAACCGCGCATGCTGCGCTGTGCCCGCGCGCATTTCTCTCCC	131
<i>Db</i>	1260	Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y:::	1319
<i>Qy</i>	132	CGTGCTTCCTTCTCTGGGCTTCGTGCTCTCTCGCTCGCATGGCGGACGGCGCT	191
<i>Db</i>	1320	Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y:::	1379
<i>Qy</i>	192	GCAACACAGGCGCCTTCCCTCGATACGGTCACTTCTACAAGGTCAATTCCCAAAGC	251
<i>Db</i>	1380	Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y::: Y:::	1439
<i>Qy</i>	252	AAGTTGGTCT	261
<i>Db</i>	1440	AAATTTCTCT	1449

```

RESULT 10
US-09-949-016-17117/c
; Sequence 17117, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17117
; LENGTH: 43267
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(43267)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17117

```

Query Match	3.5%	Score 42.2;	DB 4;	Length 43267;
Best Local Similarity	53.1%;	Pred. No. 0.34;		
Matches 112;	Conservative 0;	Mismatches 98;	Indels 1;	Gaps 1;
QY	5	GGGTGAGTGGCCCCCAGTCACGTGACCGCTGACTGGGGCGGTTCCTCCACTATCGCTT	64	
Db	23165	GCCTGCCCCACATGTCCTCCGGGCTGCCCCACATGCTCGGGCTGCCCCACATGTCGCCA	23106	
QY	65	ACCTACTCTCCCTGTGCGAAACCCGCGCATATGGCTGGCGTGTGCTCCCGCGCGCGCATTT	124	

[illegible]

```

RESULT 11
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(289)
; OTHER INFORMATION: n = A,T,C or G
; US-09-007-005-17

```

[illegible]

RESULT 12
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihc
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS


```
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17314
; LENGTH: 86936
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(86936)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17314

Query Match      3.1%; Score 37.2; DB 4; Length 86936;
Best Local Similarity 51.2%; Pred. No. 17;
Matches 87; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

Qy 53 CCACATCGCTTACCTACCTCCCTGCGAGGACCGGGGATATGGCTGCCGCTGTGCC 112
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 95 CAACAATTGCCGTGGTCTCCCAAGTGACGGTTCTCGCCCGCCCGCTCCGCTCTCG 154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 113 CGCGCGCATTTCTCTCCCGCGCTTCCCTTCTCTGGGCTTCTGCTCTCTCCGCTC 172
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 155 CGCCCGCGCCCTCCCGGCTCTCCCGGCCCTCCCGGCTCTCCCGGCCCTCCCGGCTC 214
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 173 CGCATGGGCGGAGCGGCTTGACACCAAGGCGCCCTTCCCTGGATACGG 222
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 215 CTCCCGCGCCCTCCCGGCCCTTCCCGGCCCTCAACCTAGTAGGG 264
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: July 3, 2005, 16:59:32
Job time : 193.034 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 703.092 Seconds

(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-7

Perfect score: 1205

Sequence: 1 GCGCGGAGTCCGCCCC.....TCAGTGTTCGTTAAGTGG 1205

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1201	99.7	1205	8	US-08-731-499-7
2	1062.8	88.2	1334	9	US-09-925-301-566
3	803.8	66.7	963	15	US-10-205-823-43
4	803.8	66.7	963	21	US-10-956-157-1670
5	517.2	42.9	4529	17	US-10-191-803-273
6	517.2	42.9	4529	18	US-10-152-319A-2052
7	515.8	42.8	600	21	US-10-956-157-6905

8	496.6	41.2	1139	17	US-10-388-934-523	Sequence 523, Appl
9	404.8	33.6	438	17	US-10-242-535A-37352	Sequence 37352, A
10	404.8	33.6	438	18	US-10-085-783A-37352	Sequence 37352, A
11	328.8	27.3	395	14	US-10-066-543-1394	Sequence 1394, Ap
12	321	26.6	405	10	US-09-918-995-5850	Sequence 5850, Ap
13	279.2	23.2	450	9	US-09-864-761-2375	Sequence 2375, Ap
14	271.8	22.6	345	10	US-09-918-995-18955	Sequence 18955, A
15	252.4	20.9	255	17	US-10-242-535A-31517	Sequence 31517, A
16	252.4	20.9	255	18	US-10-085-783A-31517	Sequence 31517, A
17	252	20.9	255	17	US-10-242-535A-23837	Sequence 23837, A
18	225.4	18.7	281	18	US-10-085-783A-23837	Sequence 23837, A
19	225.4	18.7	281	18	US-10-424-599-16936	Sequence 16936, A
20	223.8	18.6	248	20	US-10-424-599-16936	Sequence 16936, A
21	220.2	18.3	360	21	US-10-425-115-120840	Sequence 120840,
22	218.8	18.2	270	20	US-10-425-115-59573	Sequence 59573, A
23	213.8	17.7	383	9	US-09-960-352-14767	Sequence 14767, A
24	186.8	15.5	197	17	US-10-242-535A-19904	Sequence 19904, A
25	186.8	15.5	197	18	US-10-085-783A-19904	Sequence 19904, A
26	165	13.7	165	14	US-10-066-543-2866	Sequence 2866, Ap
27	151.4	12.6	415	9	US-09-960-352-12723	Sequence 12723, A
28	140	11.6	210	17	US-10-242-535A-4438	Sequence 4438, Ap
29	140	11.6	210	18	US-10-085-783A-4438	Sequence 4438, Ap
30	99.6	8.3	119	9	US-09-864-761-19109	Sequence 19109, A
31	90	7.5	125	17	US-10-242-535A-17878	Sequence 17878, A
32	90	7.5	125	18	US-10-085-783A-17878	Sequence 17878, A
33	50	4.1	50	17	US-10-131-827-2967	Sequence 2967, Ap
34	50	4.1	50	17	US-10-131-827-3824	Sequence 3824, Ap
35	46.4	3.9	60	10	US-09-908-375-7933	Sequence 7933, Ap
36	42.4	3.5	999	14	US-10-184-644-434	Sequence 434, App
37	42.4	3.5	999	14	US-10-184-644-434	Sequence 434, App
38	41.8	3.5	220895	19	US-10-775-169-88	Sequence 88, Appl
39	41.6	3.5	594	14	US-10-123-155-10	Sequence 10, Appl
40	41.6	3.5	594	15	US-10-146-731-10	Sequence 10, Appl
41	41.6	3.5	594	15	US-10-140-472-10	Sequence 10, Appl
42	41.6	3.5	594	15	US-10-141-761-10	Sequence 10, Appl
43	41.6	3.5	594	16	US-10-142-885-10	Sequence 10, Appl
44	41.6	3.5	594	16	US-10-158-790-10	Sequence 10, Appl
45	41.6	3.5	594	17	US-10-137-871-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1
US-08-731-499-7
; Sequence 7, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1205 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 1..1205
; LOCATION: 1..1205
; OTHER INFORMATION: /note="cDNA clone lb4 for a serine
; OTHER INFORMATION: threonine kinase"
US-08-731-499-7

Query Match 99.7%; Score 1201; DB 8; Length 1205;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1205; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGGCGGTGAGTCGCGCCCGCCAGTCAGTCAGCGCTGACTCGGGGGGTTCTCCACTATC 60
DB 1 GCGGCGGTGAGTCGCGCCCGCCAGTCAGTCAGCGCTGACTCGGGGGGTTCTCCACTATC 60
QY 61 GCTTACCTACCTCCCTCTGCGAGGACCGCGGATATGCTCGCGTGGCCCGCGCGC 120
DB 61 GCTTACCTACCTCCCTCTGCGAGGACCGCGGATATGCTCGCGTGGCCCGCGCGC 120
QY 121 ATTTCTCTCCCGCTGCTTCCCTCTCTCGGGCTTCTGCTCTCTCGCTCCGATGGC 180
DB 121 ATTTCTCTCCCGCTGCTTCCCTCTCTCGGGCTTCTGCTCTCTCGCTCCGATGGC 180
QY 181 GCGAGCGCGCTGCGACCAAGGCGCCCTTCCCTGGATACGGTGACTTCTACAGGTCA 240
DB 181 GCGAGCGCGCTGCGACCAAGGCGCCCTTCCCTGGATACGGTGACTTCTACAGGTCA 240
QY 241 TTCCCAAGCAAGTTCTGCTGGTGAAGTTCGACACCCAGTACCCCTACGGTGAGAACA 300
DB 241 TTCCCAAGCAAGTTCTGCTGGTGAAGTTCGACACCCAGTACCCCTACGGTGAGAACA 300
QY 301 GGATGAGTTCAAGCGTCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAG 360
DB 301 GGATGAGTTCAAGCGTCTTCTGAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAG 360
QY 361 GTGGGATCTCAGATTATGTGACAAAGCTGAACATGAGCTGAGTGAGAAATACAAAGTGG 420
DB 361 GTGGGATCTCAGATTATGTGACAAAGCTGAACATGAGCTGAGTGAGAAATACAAAGTGG 420
QY 421 ACAAGAGAGTACCCATCTTCTACTCTCTCGGGATGGGACTTTGAGAACCCAGTCCC 480
DB 421 ACAAGAGAGTACCCATCTTCTACTCTCTCGGGATGGGACTTTGAGAACCCAGTCCC 480
QY 481 ATACACTGGGCGAGTTAGTGTGGACCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAC 540
DB 481 ATACACTGGGCGAGTTAGTGTGGACCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAC 540
QY 541 CTAGGTATGCTGTGTCCTGCTGTATACGACGCCCTGGCGGGGAGTTCTATCAGGGCCT 600
DB 541 CTAGGTATGCTGTGTCCTGCTGTATACGACGCCCTGGCGGGGAGTTCTATCAGGGCCT 600
QY 601 CTGTTGTGAGGCGCCGAGGCCCTCTTGAAGCAGGGCAAGATACCTCTCAAGTGTAA 660
DB 601 CTGTTGTGAGGCGCCGAGGCCCTCTTGAAGCAGGGCAAGATACCTCTCAAGTGTAA 660
QY 661 GGAGACTCAGAAGAGTGGGCGCGAGCAATACCTGAAAGATCATGGGGAAGATCTTAGACCA 720
```

```

; DB 661 GGAGACTCAGAAGAGTGGGCGCGAGCAATACCTGAAGATCATGGGAAGATCTTAGACCAA 720
; QY 721 GGGGAGCACTTCCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAGAACAGA 780
; DB 721 GGGGAGCACTTCCAGCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAAGAACAGA 780
; QY 781 TGAGTGACGCGCAGAAAGAGGAGCTCCAGAAAGAGCTTAAACATCTCTGACTGCTTCCAGAA 840
; DB 781 TGAGTGACGCGCAGAAAGAGGAGCTCCAGAAAGAGCTTAAACATCTCTGACTGCTTCCAGAA 840
; QY 841 GAAGGGGCGCGAGAAAGAGAGCTGTAAAAAGGCTGTGTGATTTTCCAGGGTTTGGTG 900
; DB 841 GAAGGGGCGCGAGAAAGAGAGCTGTAAAAAGGCTGTGTGATTTTCCAGGGTTTGGTG 900
; QY 901 GGGGTAGGAGGGGAGGANNAGTTAACTCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGY 960
; DB 901 GGGGTAGGAGGGGAGGANNAGTTAACTCTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGY 960
; QY 961 MSKGGGAAAGWGGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCCCTGGACATTGATGC 1020
; DB 961 MSKGGGAAAGWGGTACTTAACCCACGATTTCTGAGCCCTGAGTATGCCCTGGACATTGATGC 1020
; QY 1021 TAAATGACCATGCTTGGGATGTCTCTAGCTGTCTGGGATAGCTGGAGACATTACTCA 1080
; DB 1021 TAAATGACCATGCTTGGGATGTCTCTAGCTGTCTGGGATAGCTGGAGACATTACTCA 1080
; QY 1081 GGTGGCTGGTGAATGACACTCAGAGGAGTAAATGAGTGTCTATAGAGGAGAGAGAGTGT 1140
; DB 1081 GGTGGCTGGTGAATGACACTCAGAGGAGTAAATGAGTGTCTATAGAGGAGAGAGAGTGT 1140
; QY 1141 ACTGCCAGGCTTTTGACAGATGTAATCTCAATTCAATTTAAAGTTTCAAGTGTGGTTA 1200
; DB 1141 ACTGCCAGGCTTTTGACAGATGTAATCTCAATTCAATTTAAAGTTTCAAGTGTGGTTA 1200
; QY 1201 AGTGG 1205
; DB 1201 AGTGG 1205

RESULT 2
US-09-925-301-566
; Sequence 566, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 566
; LENGTH: 1334
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1253)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1307)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1309)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1312)
; OTHER INFORMATION: n equals a,t,g, or c
```


QY 264 TGAAGTTGACACCCAGTACCCCTACGGTGAAGCAGGATGAGTTCAAGGCTCTT-CTG 322
DB 184 TGAAGTTGACACCCAGTACCCCTACGGTGAAGCAGGATGAGTTCAAGGCTCTTGTG 243
QY 323 AAAAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT-CTG 381
DB 244 AAAAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTATG 303
QY 382 ACAAGCTGAACATGAGCTGAGTGAAGAAATACAAGCTGGAACAGAGAGCTACCCA-TCT 440
DB 304 ACAAGCTGAACATGAGCTGAGTGAAGAAATACAAGCTGGAACAGAGAGCTACCCA 363
QY 441 TCTACTCTTCCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGAGGCGAGTT-AGG 499
DB 364 TCTACTCTTCCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGAGGCGAGTTA 423
QY 500 TTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGGTATGCTGG-TC 558
DB 424 TTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGGTATGCTGGT 483
QY 559 TGCCCTGTATACGACGCCCTGGCCGGGAGTTTCAAGGCTGAGTGAAGAGCTCAGAA 617
DB 484 TGCCCTGTATACGACGCCCTGGCCGGGAGTTTCAAGGCTGAGTGAAGAGCTCAGAA 543
QY 618 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAA 676
DB 544 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAA 603
QY 677 GGGCCGAGCAATACCTGAAGATCATGGGAGATCTTGAACAAAGGAGCACTT-CA 735
DB 604 GGGCCGAGCAATACCTGAAGATCATGGGAGATCTTGAACAAAGGAGCACTTCC 663
QY 736 CATCAGAGTGAACGAGTCCGCGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 794
DB 664 CATCAGAGTGAACGAGTCCGCGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 723
QY 795 AGGAGGAGTCCAGAGAGCTTAAACATCTTCACTGCTTCCAGAGAGGAGGAGG 854
DB 724 AGGAGGAGTCCAGAGAGCTTAAACATCTTCACTGCTTCCAGAGAGGAGGAGG 783
QY 855 AAGAGGAGTGAAGAGGCTCTGTGATTTTCCAGGTTTGGTGGGAGTGAAGAGG 914
DB 784 AAGAGGAGTGAAGAGGCTCTGTGATTTTCCAGGTTTGGTGGGAGTGAAGAGG 843
QY 915 ANAGTTAACTCTGGCTGTGANTCCCTTGTGAATATAAGGGGAGGAGGAGG 974
DB 844 AGAGTTAACTCTGGCTGTGANTCCCTTGTGAATATAAGGGGAGTGAAGAGGAGG 903

RESULT 4

US-10-956-157-1670
; Sequence 1670, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1670
; LENGTH: 963
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-1670

Query Match
Best Local Similarity 66.7%; Score 803.8; DB 21; Length 963;
Matches 934; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
QY 87 CCGCGGATATGGCTGCGGCTGTGCGCGCGCGGATTTCTCTCCCGCTGTT-CCCTTC 145
DB 4 CCGCGGATATGGCTGCGGCTGTGCGCGCGCGGATTTCTCTCCCGCTGTTCCCTTC 63
QY 146 TCCTGGGCTTCT 204
DB 64 TCCTGGGCTTCT 123
QY 205 CCCTTCCCTCGATACCGGTCACTTTTCTACAAGGTCACTTCCCAAAAGCAAGTTCTG 263
DB 124 CCCTTCCCTCGATACCGGTCACTTTTCTACAAGGTCACTTCCCAAAAGCAAGTTCTG 183
QY 264 TGAAGTTGACACCCAGTACCCCTTACCGGTGAAGAGAGGATGAGTTCAAGGCTTT-CTG 322
DB 184 TGAAGTTGACACCCAGTACCCCTTACCGGTGAAGAGAGGATGAGTTCAAGGCTTT 243
QY 323 AAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT-CTG 381
DB 244 AAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTAT 303
QY 382 ACAAGCTGAACATGAGCTGAGTGAAGAAATACAAGCTGGAACAAAGAGAGCTACCCA-TCT 440
DB 304 ACAAGCTGAACATGAGCTGAGTGAAGAAATACAAGCTGGAACAAAGAGAGCTACCCA 363
QY 441 TCTACTCTTCCGGATGGGAGCTTTTGAAGAACCCAGTCCCATACACTGAGGAGCTT-AGG 499
DB 364 TCTACTCTTCCGGATGGGAGCTTTTGAAGAACCCAGTCCCATACACTGAGGAGCTT 423
QY 500 TTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGGTATGCTGG-TC 558
DB 424 TTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGTCTACCTAGGTATGCTGGT 483
QY 559 TGCCCTGTATACGACGCCCTGGCCGGGAGTTTCACTCAGGGCTCTCTGGTGGAGG-CC 617
DB 484 TGCCCTGTATACGACGCCCTGGCCGGGAGTTTCACTCAGGGCTCTCTGGTGGAGG 543
QY 618 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAA 676
DB 544 AGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAA 603
QY 677 GGGCCGAGCAATACCTGAAGATCATGGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 735
DB 604 GGGCCGAGCAATACCTGAAGATCATGGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 663
QY 736 CATCAGAGTGAACGAGTCCGCGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 794
DB 664 CATCAGAGTGAACGAGTCCGCGGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAG 723
QY 795 AGGAGGAGTCCAGAGAGCTTAAACATCTTCACTGCTTCCAGAGAGGAGGAGG 854
DB 724 AGGAGGAGTCCAGAGAGCTTAAACATCTTCACTGCTTCCAGAGAGGAGGAGG 783
QY 855 AAGAGGAGTGAAGAGGCTCTGTGATTTTCCAGGTTTGGTGGGAGTGAAGAGG 914
DB 784 AAGAGGAGTGAAGAGGCTCTGTGATTTTCCAGGTTTGGTGGGAGTGAAGAGG 843
QY 915 ANAGTTAACTCTGGCTGTGANTCCCTTGTGAATATAAGGGGAGGAGGAGG 974
DB 844 AGAGTTAACTCTGGCTGTGANTCCCTTGTGAATATAAGGGGAGTGAAGAGGAGG 903
QY 975 TACTAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTTGATCTTAACATGAC 1029
DB 904 TACTAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTTGATCTTAACATGAC 958

RESULT 5

US-10-191-803-273
; Sequence 273, Application US/10191803

Publication No. US20040014040A1
GENERAL INFORMATION:
APPLICANT: MENDRICK, Donna
APPLICANT: PORTER, Mark
APPLICANT: JOHNSON, Kory
APPLICANT: HIGGS, Brandon
APPLICANT: CASTLE, Arthur
APPLICANT: ELASHOFF, Michael
TITLE OF INVENTION: Cardiotoxin Molecular Toxicology Modeling
FILE REFERENCE: 44921-5090US
CURRENT APPLICATION NUMBER: US/10/191,803
CURRENT FILING DATE: 2002-07-10
PRIOR APPLICATION NUMBER: US 60/303,819
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/305,623
PRIOR FILING DATE: 2001-07-17
PRIOR APPLICATION NUMBER: US 60/369,351
PRIOR FILING DATE: 2002-04-03
PRIOR APPLICATION NUMBER: US 60/377,611
PRIOR FILING DATE: 2002-05-06
NUMBER OF SEQ ID NOS: 1140
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 273
LENGTH: 4529
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20040014040A1 NM_053961
US-10-191-803-273

Query Match 42.9%; Score 517.2; DB 17; Length 4529;
Best Local Similarity 78.5%; Pred. No. 1.2e-158; Mismatches 201; Indels 26; Gaps 17;
Matches 832; Conservative 1;

QY 1 GCGCGCGTGAAGTCCGCCCCCAGTCACTGACCGCTGACTCGGGCGTTCCTCACTATC 60
Db 3277 GCGCGCGGAGCGGCCCTCCAGTGAAGTCAAGCGGCGCTGAGAGCGTCTCCTGTA 3335
QY 61 GCTTACCTACCTCCCTCTGAGGAAACCGGCGATATGCTGCGCTGCGCCCGCGCGC 120
Db 3336 GCTTCCCATCTCCGCGTCCCAACATCGGCGTGTGCTGCGCGCTTCTGCGCGCGT 3395
QY 121 ATTTCTCTCCCGCTGCT-TCCCTTCTCTGCGCTTCTGCTCTCTGCTCTCTGCTCT 179
Db 3396 TTGCTCTCTCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3455
QY 180 CGCAGCGCGCTGCACACAA-GGCGCGCTTCCCTGATGATGCTGCTGCTGCTGCTGCT 238
Db 3456 CGCAGCGCGCTGCACAGAGGCGCGCTTCCCTGATGATGCTGCTGCTGCTGCTGCT 3515
QY 239 CATTCCTCCAAAGCAAGTTCGTC-TGCTGAAGTTCGACACCCAGTACCCCTACGCTG 297
Db 3516 CATTCCTCCAAAGCAAGTTCGTC-TGCTGAAGTTCGACACCCAGTACCCCTATG 3575
QY 298 CGAGATGAGTTCAGCGCTCT-TCTGAAGTTCGACACCCAGTACCCCTACGCTG 356
Db 3576 CGAAGTGAAGTTCAGCGCTCT-TCTGAAGTTCGACACCCAGTACCCCTATG 3635
QY 357 AGAGTGGGATCTCAGATTAT-GTGACAAGTTCGACACCCAGTACCCCTACGCTG 415
Db 3636 AGAGTGGGATCTCAGATTAT-GTGACAAGTTCGACACCCAGTACCCCTATG 3695
QY 416 GCTGACAAAGAGCTACCCA-TCTTCTCTCTTCCGGGATGGGACTTTGAGAACCC 474
Db 3696 GCTGACAAAGAGCTACCCA-TCTTCTCTCTTCCGGGATGGGACTTTGAGAACCC 3755
QY 475 AGTCCCATACATGGGCGAGTT-AGTTGGAGCATCCAGCGCTGCTGAGAGGCGC 533
Db 3756 AGTCCCATACATGGGCGAGTT-AGTTGGAGCATCCAGCGCTGCTGAGAGGCGC 3815
QY 534 GGTCTACCTAGTATGCTGG-TGCTGCTCTGATACGACGCGCTGCGCGGCGAGTT 592
Db 3816 AGTCTATCTGGCATGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3875

QY 593 CAGGCGCTCTGCTGGAGG-CGCGCAGGCGCTCTTGAAGCAGGCGGCAAGATAAAGCTCTC 651
Db 3876 CGAGCGCTCCAGCAGAGAGGCGCGCGCAGGCGCTCTTGAAGCAGGCGGCAAGATAAAGCTCTC 3935
QY 652 AAGTGTGAAGGAGACTCAGAAG-AGTGGGCGGAGCAATACCTGAAGATCATGGGGAAGAT 710
Db 3936 AGTGTGAAGGAGAGACAGACAGAGAGTGGCCAGTCACTCTGAAGATCATGGGGAAGAT 3995
QY 711 CTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACAGGATCGCGAGGCTGATGGA 769
Db 3996 CTTGACCAAGGTGAAGACTTCCGCGCTCCGAGCTGCGCGGATCAGTAAGCTCATTTGA 4055
QY 770 GAAGAACAGATGATGATGACGCGCAGAGAGAGGCTCCAGAGAGCTTAAACATCTGACT 829
Db 4056 GAACAGATGATGATGAGG--GTAAGAGAGAGAGCTGCGAGAGGCTCAACATCTCTACC 4113
QY 830 GCCTTCCAGAAAGAGGCGCGCGAGAGAGAGCTGTAAAGAGGCTGTCTGTGATTTTCC 889
Db 4114 GCCTTCCAGAAAGAGGCGCGCGAGAGAGAGGCTGT-----GAGGCGCAGGCTGCT 4167
QY 890 AGGTTTGTGCGGCTAGGAGGAGGAGGAGTAACTGCTGCTGTGANTCCCTTGTGAA 949
Db 4168 AGGTTTGTGCGGCTAGGAGGAGGAGG--AGAGCGCGCTGCTGCTGCTGCTGCTGCTG 4223
QY 950 TATAAGGGGYSKGGGAAAGAGGAGGAGGAGTAACTGCTGCTGCTGCTGCTGCTGCTG 1009
Db 4224 TGAAGGGGCGAGTGGAGCAAGAGCTCTGAGCAGAGAGCTGTGCTGCTGCTGCTGCTG 4281
QY 1010 GACATTGATGCTAAACATGACCATGCTTGGGATGCTCTAG 1049
Db 4282 GACACTGAGGCTGTGAGACCATACCTTGGACATCTTGG 4321

RESULT 6

US-10-152-319A-2052
Sequence 2052, Application US/10152319A
Publication No. US20040072160A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Higgs, Brandon
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
TITLE OF INVENTION: Molecular Toxicology Modeling
FILE REFERENCE: 44921-5089-US
CURRENT APPLICATION NUMBER: US/10/152,319A
CURRENT FILING DATE: 2002-05-22
PRIOR APPLICATION NUMBER: US 60/292,335
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/297,523
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,925
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,810
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,807
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/303,808
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/315,047
PRIOR FILING DATE: 2001-08-28
PRIOR APPLICATION NUMBER: US 60/324,928
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: US 60/330,867
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: US 60/330,462
PRIOR FILING DATE: 2001-10-22
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2221
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2052

```
; LENGTH: 4529
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. NM_053961
US-10-152-319A-2052

Query Match      42.9%; Score 517.2; DB 18; Length 4529;
Best Local Similarity 78.5%; Pred. No. 1.2e-158;
Matches 832; Conservative 1; Mismatches 201; Indels 26; Gaps 17;

QY 1 GCGCGGTGATGTCGCCGCCCCAGTCACGTGACCGCTGACTCGGGCGTTCTTCACTATC 60
Db 3277 GCGCGGGGAGCGGCCCTCCAGTGAGCTGACCCGCCACTCGGAGCGTC-CTGCTGTA 3335

QY 61 GCTTACCTACCTCCTCTCGAGGACCGCGGATATGCTGCGCTGTCGCCCGCGCGC 120
Db 3336 GCTTGCCCATCTCCCGCTCCCAACAATCGGCGTGATGCTGCGCGCTTCTGGGCGCGT 3395

QY 121 ATTCTCTCCCGCTGCT-TCCCTTCTCTCGGCTTCTGCTCTCTCTCGCTCCGATGG 179
Db 3396 TTGCTCTCCCACTGCTGCTGCTTCTCTTGCCCTGCTGCTCTCTCTGCTCTCATGG 3455

QY 180 CGGCGAGGCGCTGCGACCAA-GGCGCCCTTCCCTCGATAGCGTCACTTTCTACAGGT 238
Db 3456 CGCCAGCGGCTGCGACAGAGGGCGCCCTTCCCTTGACACAGTCACTTTCTACAGGT 3515

QY 239 CATTCCTCCAAAGCAAGTTGCTC-TGCTGAAGTTCCGACCCAGTACCCCTACGGTGAA 297
Db 3516 CATTCCTCCAAAGCAAGTTGCTTCTTGGAAGTTCCGACCCAGTACCCCTATGAGAGAA 3575

QY 298 CGAGGATGAGTTCAAGCGTCT-TCTGAAACTCGGCTTCCAGCGATGATCTCTGGTGGC 356
Db 3576 GCAAGATGAGTTAAGCGTCTGGCTGAGACTCAGCTCCAGCGATGATCTCTGGTGGC 3635

QY 357 AGAGTGGGATCTCAGATTAT-GTGACAGCTGAACATGGAGCTGAGTGAGAAATACAA 415
Db 3636 AGAGTGGGATCTCAGACTATGCTGACAACTGAACATGGAGCTGAGTGAGAAATACAA 3695

QY 416 GCTGACAAAGAGACTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTGACAAACC 474
Db 3696 GCTGACAAAGAGACTACCCAGTCTTCTACCTCTTCCGGGATGGGACTTTGAGAAATCC 3755

QY 475 AGTCCCATACATCGGGGAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 533
Db 3756 TGTCCTATACAGGGGAGTTAAAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGG 3815

QY 534 GGTCTACCTAGTATGCTCGG-TGCTGCTGTATACGACGCTTGGCGGGGAGTTTCAAT 592
Db 3816 AGTCTATCTGGCATGCTTGGATGCTGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 3875

QY 593 CAGGCGCTCTGCTGGAGG-CGCGCGGCGCTTGAAGCGGGGCAAGATAACCTCTC 651
Db 3876 CGAGGCTTCCAGCAGAGAGCCCGCAGCCATCTCTGAACAGGGGCGAGATGGGCTCTC 3935

QY 652 AAGTGTGAAGAGACTCAGAAG-AGTGGGCCGAGCAATACCTGAAGATCATGGGAAGAT 710
Db 3936 AGGTGTGAAGAGACAGACAAGATGGGCCAGTCACTGATCACTGAAGATCATGGGAAGAT 3995

QY 711 CTTAGACCAAGGGGAGCACTT-CCAGCATCAGAGATGACAGGATCGCCAGGCTGATTGA 769
Db 3996 CTTGACCAAGGTGAAGACTTCCCGGCTTCCAGGCTGGCCCGGATCAGTAAGCTCATTTGA 4055

QY 770 GAAGACAGATGAGTGAGG--GTAAAGAAAGAGAGCTCCAGAGAGCTTAAACATCTCTGCT 829
Db 4056 GAAACAGATGAGTGAGG--GTAAAGAAAGAGAGCTCCAGAGAGCTTAAACATCTCTCACC 4113

QY 830 GCCTTCCAGAAAGGGGCCGAGAAAGAGAGCTGTAAAAAGGCTGTCTGTGATTTTCC 889
Db 4114 GCCTTCCGAAAGAGCCCGGAGAGAGAGCTGT-----GAGGGCACCAGGCTGCT 4167

QY 890 AGGTTTGTGGGGTAGGGAGGGGAGGAGTTAACTCTGCTGCTGTGANTCCCTTGTGGAA 949
Db 949 AGGTTTGTGGGGTAGGGAGGGGAGGAGTTAACTCTGCTGCTGTGANTCCCTTGTGGAA 949
```

```
Db 4168 AGGTTTGTGAGGGCGGGGAGGG--AGAGCCGCTGCTGACTGTGGGGCCCTGTGGG-- 4223
QY 950 TATAAGGGGGYMSKGGGAAAGMGGTACTATAACCCACGATTTCTGAGCCCTGAGTATGCTG 1009
Db 4224 TGGAGGGGCGAGTGGAGCAAGCAGTCTTGAGCCAGAGACCTGTGCCCCGAG--TGCTG 4281
QY 1010 GACATTGATGCTAAACAGCACCATCTCTTGGGATGCTCTAG 1049
Db 4282 GACACTGGGCTGCTGAGACCATACCTCTGGACATCCTTGG 4321

RESULT 7
US-10-956-157-6905
; Sequence 6905, Application US/10956157
; Publication No. US20050118625A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William
; TITLE OF INVENTION: NUCLEIC ACID ARRAYS FOR DETECTING GENE EXPRESSION ASSOCIATED WITH
; FILE OF INVENTION: HUMAN OSTEOARTHRITIS AND HUMAN PROTEASES
; FILE REFERENCE: 031896-043000 (AM 101081)
; CURRENT APPLICATION NUMBER: US/10/956,157
; CURRENT FILING DATE: 2004-10-04
; NUMBER OF SEQ ID NOS: 319805
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6905
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-956-157-6905

Query Match      42.8%; Score 515.8; DB 21; Length 600;
Best Local Similarity 97.5%; Pred. No. 1.3e-158;
Matches 580; Conservative 5; Mismatches 4; Indels 6; Gaps 6;

QY 441 TCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTT-AGG 499
Db 1 TCTACTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGCGAGTTAAGG 60

QY 500 TTGAGGCCATCCAGCGCTGGCTGAAGGGCAAGGGTCTACTAGGTATGCTGG-TGCC 558
Db 61 TTGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGTCTACTAGGTATGCTGGTTGCC 120

QY 559 TGCTGTATACGACGCCCTGGCCGGGGAGTTTCATCAGGGCTCTGGTGTGAGG-CCGCC 617
Db 121 TGCTGTATACGACGCCCTGGCCGGGGAGTTTCATCAGGGCTCTGGTGTGAGGCCGCC 180

QY 618 AGGCCCTTTCAGAGCGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAG-AGT 676
Db 181 AGGCCCTTTCAGAGCGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAGAAAGT 240

QY 677 GGGCGAGCAATACCTGAAGATCATGGGAAAGATCTTAGACCAAGGGGAGCACTT-CCAG 735
Db 241 GGGCGAGCAATACCTGAAGATCATGGGAAAGATCTTAGACCAAGGGGAGCACTTCCAG 300

QY 736 CATCAGAGATGACCGATCCGAGTGTGTTGAGAAAGAAACAAGATGATGACGCG-AGA 794
Db 301 CATCAGAGATGACCGATCCGAGTGTGTTGAGAAAGAAACAAGATGATGACGCGCAAGA 360

QY 795 AGGAGGAGCTCCAGAGAGCTTAAACATCTCTGAGTCCCTCCAGAGAGGGGCGCGAGA 854
Db 361 AGGAGGAGCTCCAGAGAGCTTAAACATCTCTGAGTCCCTCCAGAGAGGGGCGCGAGA 420

QY 855 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCAGGGTTTGGTGGGGTAGGGAGGGG 914
Db 421 AAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTTCAGGGTTTGGTGGGGTAGGGAGGGG 480

QY 915 ANAGTTAACTCTGCTGCTGANTCCCTTGTGGAATATAGGGGGYMSKGGGAAAGGG 974
Db 481 AGAGTTAACTCTGCTGCTGANTCCCTTGTGGAATATAGGGGGTAGGGGAAAGCTGG 540

QY 975 TACTAACCCAGATTCAGCCCTGAGTATGCTGGACATTGATCTCAATGAC 1029
Db 1029 TACTAACCCAGATTCAGCCCTGAGTATGCTGGACATTGATCTCAATGAC 1029
```

Db 541 TACTAACCCACGATCTGAGCCCTGAGTATGCTGACATTTGATGCTAACATGAC 595

RESULT 8
US-10-388-934-523
; Sequence 523, Application US/10388934
; Publication No. US20040005547A1
; GENERAL INFORMATION:
; APPLICANT: Boess, Franziska
; APPLICANT: Suter-Dick, Laura
; APPLICANT: Wolf, Detlef
; TITLE OF INVENTION: BIOMARKERS AND EXPRESSION PROFILES FOR TOXICOLOGY
; FILE REFERENCE: 21199
; CURRENT APPLICATION NUMBER: US/10/388,934
; CURRENT FILING DATE: 2003-03-14
; PRIOR APPLICATION NUMBER: 02005336.9
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 02015657.6
; PRIOR FILING DATE: 2002-07-17
; NUMBER OF SEQ ID NOS: 862
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 523
; TYPE: DNA
; ORGANISM: Rattus norvegicus (No. US20040005547Alway rat)
US-10-388-934-523

Query Match 41.2%; Score 496.6; DB 17; Length 1139;
Best Local Similarity 78.8%; Pred. No. 3.8e-152;
Matches 792; Conservative 1; Mismatches 187; Indels 25; Gaps 16;

QY 56 CTATGCTTACTACTCTCCCTCTGCAGGACCCGGGATATGGCTGCGCTGTCGCGCCG 115
DB 5 CTGTAGCTTGCCTCTCCCTCTCCCAATCCGCGGTGATGGCTGCGCGCTTCTCTGC 64

QY 116 GCGCGATTTCTCCCGCTGCT-TCCCTTCTCTGGGCTTCTGCTCTCTCCGCTCG 174
DB 65 GCGGTTTCTCTCCCACTGCTGCTGCTTCTCTGGCTGCTGCTCTCTCTCTCTCT 124

QY 175 CATGCGGCGAGCGGCTGCACACAA-GGCGCCCTTCCCTGGATACCGTCACTTTCTAC 233
DB 125 CATGCGGCGAGCGGCTGCACACGAAGGCGCGCTTCCCTGGACACAGTCACTTTCTAC 184

QY 234 AGGTTCATTTCCCAAGCAAGTTGCTC-TGGTGAAGTTGCACACCCAGTACCCCTACGGT 292
DB 185 AAGGTCATTTCCCAAGCAAGTTGCTC-TGGTGAAGTTGCACACCCAGTACCCCTATGGA 244

QY 293 GAGAAGCAGATGAGTTCAAGGCTCT-TCTGAAACTCGGCTTCCAGCGATGATCTCTTG 351
DB 245 GAGAAGCAGATGAGTTAAGGCTCTGGCTGAGACTAGCTTCCAGCGATGATCTCTTG 304

QY 352 GTGGCAGAGTGGGATCTCAGATTAT-GTGAACAAGCTGAAATGAGCTGAGTGAGAAA 410
DB 305 GTGGCAGAGTGGGATCTCAGACTATGCTGACACAGCTGAAATGAGCTGAGTGAAG 364

QY 411 TACAAGCTGGAACAAGAGAGTACCCA-TCTTCTACTCTTCCGGATGGGAGCTTTGAG 469
DB 365 TACAAGCTGGAACAAGAGAGTACCCA-TCTTCTACTCTTCCGGATGGGAGCTTTGAG 424

QY 470 AACCCAGTCCCATACACTGGGCGAGTT-AGGTTGAGCCATCCAGGCTGGCTGAAGGG 528
DB 425 AATCTGTCCCATACAGCGGGGAGTTAAAGTTGAGCCATCCAGGCTGGCTGAAGGG 484

QY 529 CAAGGGGTCTACTAGGTATGCTGG-TGCTGCTGCTGTATACGACGCGCTTCCCGGGAG 587
DB 485 CAGGAGTCTATCTGGGATGCTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544

QY 588 TTCATCAGGCTCTGCTGCTGAGG-CGCGCAGGCTCTTGAAGCAGGGGCAAGATAAC 646
DB 545 TTCATCAGGCTCTCAGCAGAGAGGCGCCGAGGCGATCTCTGAAACAGGGGCGAGTGGC 604

QY 647 CTCTCAAGTGTGAAGAGCTCAGAG-AGTGGGCGGAGCATACCTGAAGATCATGGG 705

Db 605 CTCTCAGGTGTGAAGGACACACAAGAGTGGCCAGTCACTACCTGAAGATCATGGGG 664

QY 706 AAGATCTTTAGACCAAGGGAGCATT-CCAGCATCAGAGATACACGATCGCCAGGCTG 764
DB 665 AAGATCTTTGACCAAGGTGAAGACTTCCCGGCTCCGAGCTGGCCGATCAGTAAGCTC 724

QY 765 ATTGAAGAACAACAGATGAGTGACCGCAGAGGAGGAGCTCCAGAGAGCTTTAAACATCC 824
DB 725 ATTGAAGAACAAGATGAGTGAGG--GTAAGAAGAAGAGCTCGAGAGGAGCTCAACATCC 782

QY 825 TGACTGCTTCCAGAAAGGGGCGGAGAAAGAGAGCTGTAAAGAGCTGTCTGTGAT 884
DB 783 TCACCGCTTCCGCAAGAAAGCGCGGAGAGGAGCTGT-----GAGGCGACCAAG 836

QY 885 TTTCCAGGTTTGTGGGGTAGGAGGGGAGNAGTTAACCTGCTGCTGTGTGANTCCCTTG 944
DB 837 CTGCTAGGTTTGTCTGAGGCGGGGAGGG--AGAGCGGCTGCTGACTGTGGGGCCCTGT 894

QY 945 TGGATATAAGGGGYSKSGGAAAGGTTACTAACCCAGATTTCTGAGCCCTGAGTAT 1004
DB 895 GGG--TGGAGGGGCGAGTGGAGCAAGAGCTCTGAGCCAGAGACTGTGCCCCGAG--T 950

QY 1005 GCCTGACATTTGATCTAACATGACCATGCTTGGGATGTCTCTAG 1049
DB 951 GCCTGACATGAGGCTGCTGAGACCATACCTTGACATCTCTGG 995

RESULT 9
US-10-242-535A-37352
; Sequence 37352, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 37352
; LENGTH: 438
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (24)..(25)
; OTHER INFORMATION: n is a, c, g, or t
US-10-242-535A-37352

Query Match 33.6%; Score 404.8; DB 17; Length 438;
Best Local Similarity 97.0%; Pred. No. 4e-122;
Matches 425; Conservative 5; Mismatches 6; Indels 2; Gaps 2;

QY 691 CTGAAGATCATGGGAGAGATCTTAGACCAAGGGAGGAGCTT-CCAGCATCAGAGATGACA 749
DB 1 CTGAAGATCATGGGAGAGATCTTNNACCAAGGGAGGAGCTTCCAGCATCAGAGATGACA 60

QY 750 CGATCGCCAGCTGATTGAGAAGAAAGATGATGAC-GGCAGAAGAGAGAGCTCCAG 808
DB 61 CGATCGCCAGCTGATTGAGAAGAAAGATGATGACGCGGAAGAGAGAGCTCCAG 120

QY 809 AAGAGCTTAAACATCTGACTGCTTCCAGAGAGGGGGCGAGAAAGAGAGCTGTAA 868
DB 121 AAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGGCGAGAAAGAGAGCTGTAA 180

Db 301 AGCTCCANAGAGCTTAACATCTGACTGCTTCCAGAGAGGGGGCCGNAAGAGGG 360
QY 861 AGCTGTAAAGGCTGTCTGATTTTCCAGGTT 895
Db 361 AGCTGTAAAGGCTGTCTGATTTTCCAGGTT 395

RESULT 12

US-09-918-995-5850
; Sequence 5850, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5850
; LENGTH: 405
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-5850

Query Match 26.6%; Score 321; DB 10; Length 405;
Best Local Similarity 98.5%; Pred. No. 1.5e-94;
Matches 387; Conservative 0; Mismatches 0; Indels 6; Gaps 6;
QY 88 CGGCGATATGGTCGGCTGTGTCGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCTTCT 146
Db 13 CGGCGATATGGTCGGCTGTGTCGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTTCT 72
QY 147 CTTGGCTTCTGCTCTCTCCGCTCCGATCGGCGCGCGCGCTTGCACACAA-GGCGC 205
Db 73 CTTGGCTTCTGCTCTCTCCGCTCCGATCGGCGCGCGCGCTTGCACACAAAGGGGC 132
QY 206 CTTCCCTTGGATAGCTGCTTCTTACAGGTCTTCCAAAGCAAGTTCGTC-TGCT 264
Db 133 CTTCCCTTGGATAGCTGCTTCTTACAGGTCTTCCAAAGCAAGTTCGTC-TGCT 192

QY 265 GAAGTTGCACACCCAGTACCCCTACGGTGAGAGCAGGATGAGTTCAAGCGTCTT-CTGA 323
Db 193 GAAGTTGCACACCCAGTACCCCTACGGTGAGAGCAGGATGAGTTCAAGCGTCTTCTGA 252
QY 324 AAATCTGGCTTCCAGGATGATCTCTTGGTGACAGGTTGGGATCTCAGATTAT-GTGA 382
Db 253 AAATCTGGCTTCCAGGATGATCTCTTGGTGACAGGTTGGGATCTCAGATTATGTGA 312
QY 383 CAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGCAAGAGAGCTACCCA-TCCT 441
Db 313 CAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGCAAGAGAGCTACCCAATCTT 372

QY 442 CTACCTCTTCCGGGATGGGATTTTGAACACC 474

Db 373 CTACCTCTTCCGGGATGGGATTTTGAACACC 405

RESULT 13

US-09-864-761-2375
; Sequence 2375, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/006666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/006661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2375
; LENGTH: 450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL109984.5
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 12
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 14
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
US-09-864-761-2375

Query Match 23.2%; Score 279.2; DB 9; Length 450;
Best Local Similarity 84.0%; Pred. No. 9.1e-81;
Matches 377; Conservative 0; Mismatches 59; Indels 13; Gaps 5;

QY 488 GGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTA 547
Db 5 GGGGCGAGTTAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTAGCTAGCTA 64
QY 548 T-----GCCTGCTGCTGCTGTATACGACGCCCTGGCGGGGAGTTCTATCAGGGCC 599
Db 65 TGCCAGGTGGGCTGGTGGCCAGCTGCATAGGACACCTGGCC-AGGAGTTCTATCAGGGCC 123
QY 600 TCTGGTGGAGG-CCGCGAGGCCCTCTTGAAGCGGGGCAAGATAACCTCTCAAGTGTG 658

Db 124 TCTGCGTGGAGCCCGCTAGGCCCTCTGGAAGCAGCGCAGGACCACTCTCAAGTGTG 183
Qy 659 AAGGAGACTCAGAAG-AGTGGCCGAGCAATACCTGGAAGATCATGGGGAAGATCTTAGAC 717
Db 184 AAGGAGACTGAGAAGAACTGGCCGAGCAATACCTGGAAGATCATGGGGAAGATCTTAGAC 243
Qy 718 CAAGGGAGCACTTCCAGCATCAGAGATGACACCGATCCGAGGCTGATTGAGAAGACA 777
Db 244 TAGGCAGAGGACTTCCAGCATAGATGACATGGATCTCCAGGCTGATTGAGAAGCAAGT 303
Qy 778 AGATGAGTGCAGCAGAGAGAGAGCTCCAGAGAGCTTAAACATCTGATGCTGCTTCCA 837
Db 304 TGAGTGACG--GGAAGAGGAGGAGCTCCAGAGAGCTTAAACATCTGATGCTTCCA 361
Qy 838 GAAGAGGGGCGCAGAGAGGAGCTCTGAAAAGGCTGTCTGATTTTCCAGGGTTG 897
Db 362 GAAGAGGGGCGTGAGAACGAGGAGTGTGAAAAGGCTCTCTATGGTTTTCCAGGGTTG 421
Qy 898 GTGGGGGTAGGAGGGGAGANAGTTAACCTG 926
Db 422 GTAGGGGTGGGAGGGAGAGTTAACCTG 450

RESULT 14

US-09-918-18955
; Sequence 18955, Application US/09918995
; Publication No. US20030073623A1

GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995

; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18955

; LENGTH: 345

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-918-995-18955

Query Match 22.6%; Score 271.8; DB 10; Length 345;
Best Local Similarity 97.9%; Pred. No. 2.2e-78;
Matches 328; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

Qy 318 TTCTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGGATCTCAGATTA 377
Db 11 TGCTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGGATCTCAAATTA 70
Qy 378 T-GTGAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGGACAAAGAGAGTACCC 436
Db 71 TGGTGACAAGCTGAACATGGAGCTGAGTGAGAAATACAACTGGGACAAAGAGAGTACCC 130
Qy 437 A-TCTTCTTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT 495
Db 131 AGTCCTTCTACCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGT 190
Qy 496 T-AGGTTGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTATGCTGG 554
Db 191 TAAGGTTGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGTCTACCTAGGTATGCTGG 250
Qy 555 -TGCTGCTGTATACGAGCCCTGGCCGGGAGTTTCATCAGGGCTCTGTGTGGAGG- 612
Db 251 TTGCTGCTGTATACGAGCCCTGGCCGGGAGTTTCATCAGGGCTCTGTGTGGAGG- 310
Qy 613 CCGCCAGGCCCTCTTGAAGCAGGGGCAAGATAACC 647
Db 311 CCGCCAGGCCCTCTTGAAGCAGGGGCAAGATAACC 345

RESULT 15

US-10-242-535A-31517
; Sequence 31517, Application US/10242535A
; Publication No. US20040013663A1

GENERAL INFORMATION:

; APPLICANT: ChondroGene Inc.

; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis

; FILE REFERENCE: 4231/2005

; CURRENT APPLICATION NUMBER: US/10/242,535A

; CURRENT FILING DATE: 2002-09-12

; PRIOR APPLICATION NUMBER: US 10/085,783

; PRIOR FILING DATE: 2002-02-28

; PRIOR APPLICATION NUMBER: US 60/305,340

; PRIOR FILING DATE: 2001-07-13

; PRIOR APPLICATION NUMBER: US 60/275,017

; PRIOR FILING DATE: 2001-03-12

; PRIOR APPLICATION NUMBER: US 60/271,955

; PRIOR FILING DATE: 2001-02-28

; NUMBER OF SEQ ID NOS: 58994

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 31517

; LENGTH: 255

; TYPE: DNA

; ORGANISM: Human

; NAME/KEY: misc feature

; LOCATION: (22)..(22)

; OTHER INFORMATION: n is a, c, g, or t

US-10-242-535A-31517

Query Match 20.9%; Score 252.4; DB 17; Length 255;

Best Local Similarity 98.0%; Pred. No. 4.6e-72;

Matches 250; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 941 CTTGTGGAATATAAGGGGYNKGGGAAAGWGTACTTAACCCAGGATTCGAGCCCTGA 1000
Db 1 CTTGTGGAATATAAGGGGNTANTGGGAAAGTGTACTTAACCCAGGATTCGAGCCCTGA 60
Qy 1001 GTATGCTGGACATTTGATGCTTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGG 1060
Db 61 GTATGCTGGACATTTGATGCTTAACATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGG 120
Qy 1061 ATAGCTGGAGCATTACTCTAGGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGCTA 1120
Db 121 ATAGCTGGAGCATTACTCTAGGTGGCTGGTGAATGACACCTCAGAAGAAATGAGTGCTA 180
Qy 1121 TAGAGAGAGAGAGAGTGTACTGCCAGGCTTTTGACAGATGTAATTCATTCAATTA 1180
Db 181 TAGAGAGAGAGAGAGTGTACTGCCAGGCTTTTGACAGATGTAATTCATTCAATTA 240
Qy 1181 AAGTTTCAGTGTGTTT 1195
Db 241 AAGTTTCAGTGTGTTT 255

Search completed: July 4, 2005, 11:51:00

Job time : 705.092 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 07:04:57 ; Search time 3810.04 Seconds
(without alignments)
12038.564 Million cell updates/sec

Title: US-08-731-499-7
Perfect score: 1205
Sequence: 1 GCGCGGTGAGTCGCGCCCC.....TCAGTGTTTGGTTAAGTGG 1205

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_gsl1.*

9: gb_gsl2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1034.8	85.9	1203	3	CR625433 full-length
2	1000.8	83.1	1152	3	CR600542 full-length
3	996.8	82.7	1375	3	CR623789 full-length
4	987.8	82.0	1139	3	CR617833 full-length
5	967.8	80.3	1119	3	CR602804 full-length
6	967.8	80.3	1119	3	CR611287 full-length
7	959.8	79.7	1111	3	CR604444 full-length
8	959.8	79.7	1111	3	CR617687 full-length
9	958.8	79.6	1110	3	CR603906 full-length
10	928.4	77.0	1137	5	CR603042 full-length
11	910.8	75.6	1062	3	CR602432 full-length
12	906.8	75.3	1058	3	CR594614 full-length
13	904.4	75.1	1098	5	CR603041 full-length
14	850.4	70.6	1093	1	CR611169 full-length
15	848	70.4	1099	1	CR582444 full-length
16	844.4	70.1	1081	5	CR343383 full-length
17	840.6	69.8	1079	5	CR405259 full-length
18	786	65.2	971	1	CR570078 full-length
19	783.8	65.0	933	1	CR543690 full-length
20	763.6	63.4	920	1	CR539757 full-length
21	763.6	63.4	944	1	CR582574 full-length
22	747.6	62.0	1088	5	CR343384 full-length
23	738.8	61.3	993	3	CR610095 full-length
24	737.8	61.2	887	1	CR553980 full-length

25	728.8	60.5	965	1	AL559014
26	721.4	59.9	941	5	CR387412
27	716.8	59.5	913	5	CR192764
28	713	59.2	838	1	AL573259
29	711.2	59.0	1005	5	CR405260
30	710.6	59.0	1088	4	CR464547
31	708	58.8	1040	5	CR332264
32	705.2	58.5	903	5	CR372694
33	703.8	58.4	869	1	AL577874
34	698	57.9	920	1	AL582586
35	698	57.9	990	5	CR367166
36	697.8	57.9	956	7	CR648391
37	696.8	57.8	883	5	CR49785
38	695.2	57.7	818	6	CR515682
39	691.8	57.4	963	5	CR367165
40	685.6	56.9	1154	4	CR455050
41	678.8	56.3	1091	5	CR920142
42	674.8	56.0	950	5	CR332263
43	673.2	55.9	1058	4	CR555831
44	672.8	55.8	1087	5	CR53366
45	670	55.6	873	6	CR385022

ALIGNMENTS

RESULT 1
CR625433
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE
ORIGIN
Query Match
Best Local Similarity
Matches 1165; Conservative
QY
DB

1203 bp mRNA linear HTC 21-JUL-2004
full-length cDNA clone CS0DK008YM02 of HeLa cells Cot 25-normalized
CR625433
CR625433.1 GI:50506240
HTC; CDSLT CDNA.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue
Genoscope.
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of invitrogen.
location/Qualifiers
1. 1203
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK008YM02"
/tissue_type="HeLa cells Cot 25-normalized"
/plasmid="pCMVSPORT_6"


```
Qy 61 GCTTACCTACCTCCCTCTGCGAGAACCGGCGATATGGCTGCGCTGTGCCCCCGCGCGC 120
Db 78 GCTTACCTACCTCCCTCTGCGAGAACCGGCGATATGGCTGCGCTGTGCCCCCGCGCGC 137

Qy 121 ATTTCTCTCCCGCTGCTT-CCCTTCTCTCGGCTTCTGCTCTCTCTCCGCTCCGCGATGG 179
Db 138 ATTTCTCTCCCGCTGCTTCCCTTCTCTCGGCTTCTGCTCTCTCTCCGCTCCGCGATGG 197

Qy 180 CGGAGCGGCTGCACACCAA-GGGGCGCTTCCCTGATAGGTCATCTTCTACAAGGT 238
Db 198 CGGAGCGGCTGCACACCAAAGGGCGCTTCCCTGATAGGTCATCTTCTACAAGGT 257

Qy 239 CATTCCCAAGCAAGTTGCTC-TGGTGAAGTTGACACCCAGTACCCTACGGTGAGAA 297
Db 258 CATTCCCAAGCAAGTTGCTCTTGGTGAAGTTGACACCCAGTACCCTACGGTGAGAA 317

Qy 298 CGAGATGAGTTCAAGCGTCTT-CTGAAAACTCGGCTTCCAGCGATGATCTTTGGTGGC 356
Db 318 GCAGATGAGTTCAAGCGTCTTCTGCTGAAAACTCGGCTTCCAGCGATGATCTTTGGTGGC 377

Qy 357 AGAGTGGGATCTCAGATTAT-GTGACAACTGACATGAGCTGAGTGAGAAATACAA 415
Db 378 AGAGTGGGATCTCAGATTATGAGTGAAGCTGAACATGAGCTGAGTGAGAAATACAA 437

Qy 416 GCTGACAAAGAGCTACCCA-TCTTCTACTCTTCCGGATGGGACTTTGAGAACC 474
Db 438 GCTGACAAAGAGCTACCCAAGTCTTCTACTCTTCCGGATGGGACTTTGAGAACC 497

Qy 475 AGTCCATACACTGGGGCAGTT-AGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAG 533
Db 498 AGTCCATACACTGGGGCAGTTAAGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCAAG 557

Qy 534 GGTCTACCTAGTATGCTGG-TGCTGCTGTATACGACGCCCTGGCGGGGAGTTTCA 592
Db 558 GGTCTACCTAGTATGCTGGTGGCTGCTGCTGTATACGACGCCCTGGCGGGGAGTTTCA 617

Qy 593 CAGGCTCTGCTGTGGAGG-CCGCGAGGCTCTTTGAAGCAGGGGCAAGATAACCTCTC 651
Db 618 CAGGCTCTGCTGTGGAGGCGCGCAGGCTCTTTGAAGCAGGGGCAAGATAACCTCTC 677

Qy 652 AAGTGTGAAGAGACTCAAGAG-AGTGGGCGAGCAATACCTGAAGATCATGGGGAAGAT 710
Db 678 AAGTGTGAAGAGACTCAAGAGAGTGGGCGAGCAATACCTGAAGATCATGGGGAAGAT 737

Qy 711 CTTAGACCAAGGAGCACTT-CCAGCATCAGAGTACAGCGATCGCCAGGCTGATTGA 769
Db 738 CTTAGACCAAGGAGCACTTCCAGCATCAGAGTACAGCGATCGCCAGGCTGATTGA 797

Qy 770 GAAGAACAGATGAGTGAC-GGCAGAGGAGGCTCCAGAGAGCTTTAAACATCCTGAC 828
Db 798 GAAGAACAGATGAGTGACGCGGAAGAGGAGGCTCCAGAGAGCTTTAAACATCCTGAC 857

Qy 829 TGCTTCCAGAGAAAGGGGCGAGAAAGAGAGCTGTAAAAGGCTGTCTGTGATTTTC 888
Db 858 TGCTTCCAGAGAAAGGGGCGAGAAAGAGAGCTGTAAAAGGCTGTCTGTGATTTTC 917

Qy 889 CAGGTTTGGTGGGATAGGAGGGGAGTTAACTCTGCTGTGANTCCCTTGTGGA 948
Db 918 CAGGTTTGGTGGGATAGGAGGGGAGTTAACTCTGCTGTGAGTCCCTTGTGGA 977

Qy 949 ATATAAGGGGYSKGGGAAAGWGTACTTAACCCAGCATCTGAGCCCTGAGTATGCT 1008
Db 978 ATATAAGGGGATAGTGGGAAAGTGTACTTAACCCAGCATCTGAGCCCTGAGTATGCT 1037

Qy 1009 GGACATTGATCAATGACCATGCTTGGGATGTCTTAGCTGTCTGGGATAGCTGG 1068
Db 1038 GGACATTGATCAATGACCATGCTTGGGATGTCTTAGCTGTCTGGGATAGCTGG 1097

Qy 1069 AGCACTTACTAGGTGGCTGGTGAATGACACCTCAGAGGAATAGTGTCTATAGAGG 1128
Db 1098 AGCACTTACTAGGTGGCTGGTGAATGACACCTCAGAGGAATAGTGTCTATAGAGG 1157
```

```
Qy 1129 AGAGGAGGTGTACTGCCAGGTCTTTGACAGATGTAAATTCATTT 1174
Db 1158 AGAGGAGGTGTACTGCCAGGTCTTTGACAGATGTAAATTCATTT 1203

RESULT 2
CR600542 1152 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CSODI006YCl2 of Placentia Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR600542
VERSION CR600542.1 GI:50481349
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1152)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1152)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : secrete@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1152
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI006YCl2"
/tissue_type="Placentia Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 83.1%; Score 1000.8; DB 3; Length 1152;
Best Local Similarity 98.2%; Pred. No. 1.2e-270;
Matches 1131; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
Qy 25 TCAGTGAACCGCTGACTCGGGCGGTTCTCCACTATCGTTACCTACCTCCCTCTGCAGGA 84
Db 1 TCAGTGAACCGCTGACTCGGGCGGTTCTCCACTATCGTTACCTACCTCCCTCTGCAGGA 60
Qy 85 ACCCGCGGATATGGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCT 143
Db 61 ACCCGCGGATATGGCTGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCT 120
Qy 144 TCTCTGGGCTTCCCTGCTCTCCGCTCCGCTCCGCGGAGCGGCTGCACACCA-GG 202
Db 121 TCTCTGGGCTTCCCTGCTCTCCGCTCCGCTCCGCGGAGCGGCTGCACACCAAGG 180
Qy 203 CGCCCTTCCCTGGATACGGTCACCTTTCTACAAGTCAATTCCTCAAAAGCAAGTTGCTC-T 261
Db 181 CGCCCTTCCCTGGATAGGTCACCTTTCTACAAGTCAATTCCTCAAAAGCAAGTTGCTCT 240
Qy 262 GGTGAAGTTGCACACCCAGTACCCCTACGGTGAAGCAGGATGAGTTCAAGCGCTT-C 320
Db 241 GGTGAAGTTGCACACCCAGTACCCCTACGGTGAAGCAGGATGAGTTCAAGCGCTTGC 300
Qy 321 TGAACACTCGGCTTCCAGCGATGATCTTGGTGGCAGAGTGGGATCTCAGATTAT-G 379
Db 301 TGAACACTCGGCTTCCAGCGATGATCTTGGTGGCAGAGTGGGATCTCAGATTATGG 360
```

```
QY 380 TCACAAGCTGAACATGGAGCTGAGTGAAGAAATACAAAGCTGGACAAAGAGAGCTACCCCA-T 438
Db 361 TCACAAGCTGAACATGGAGCTGAGTGAAGAAATACAAAGCTGGACAAAGAGAGCTACCCAGT 420
QY 439 CTTTACCTCTTCCGGATGGGAGCTTTGAGAACCCAGTCCCATACACTCTGGGCGAGTT-A 497
Db 421 CTTTACCTCTTCCGGATGGGAGCTTTGAGAACCCAGTCCCATACACTCTGGGCGAGTTAA 480
QY 498 GGTGGAGCCATCAGCGCTGCTGAAGGGGAAGGGTCTACCTAGTATGCTCTGG-TG 556
Db 481 GGTGGAGCCATCAGCGCTGCTGAAGGGGAAGGGTCTACCTAGTATGCTCTGGTTG 540
QY 557 CTTGCTGTATACGAGCCCTGCGCGGAGGTTTCATCAGGGCTCTGTTGTGGAGG-CGG 615
Db 541 CTTGCTGTATACGAGCCCTGCGCGGAGGTTTCATCAGGGCTCTGTTGTGGAGGCCG 600
QY 616 CCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAG-A 674
Db 601 CCAGGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAGAA 660
QY 675 GTGGGCGGAGCAATCCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCACTT-CC 733
Db 661 GTGGGCGGAGCAATCCTGAAGATCATGGGGAAGATCTTTAGACCAAGGGGAGCACTTCCC 720
QY 734 AGCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAGACAAAGATCAGTGAC-GGCA 792
Db 721 AGCATCAGAGATGACACGGATCGCCAGGCTGATTGAGAGACAAAGATGAGTGACGGGAA 780
QY 793 GAAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGAGGGGCGCA 852
Db 781 GAAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAGAGGGGCGCA 840
QY 853 GAAAGGAGGAGCTTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 912
Db 841 GAAAGGAGGAGCTTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGG 900
QY 913 GGAGGATTAACCTGTGCTGTGANTCCCTTGTGCAATATAGGGGYYMSKGGGAAAGW 972
Db 901 GGAGGATTAACCTGTGCTGTGANTCCCTTGTGCAATATAGGGGYYMSKGGGAAAGT 960
QY 973 GGTACTAACCCACGATTTGAGCCCTGAGTATGCTGAGACATTTGATGCTAACATGACCAT 1032
Db 961 GGTACTAACCCACGATTTGAGCCCTGAGTATGCTGAGACATTTGATGCTAACATGACCAT 1020
QY 1033 GCTTGGGATGCTCTAGTGTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGA 1092
Db 1021 GCTTGGGATGCTCTAGTGTGCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGA 1080
QY 1093 AATGACACTCAGAAGGATGAGTCTATAGAGAGAGAGAGGAGTGTACTGCCAGGTC 1152
Db 1081 AATGACACTCAGAAGGATGAGTCTATAGAGAGAGAGAGGAGTGTACTGCCAGGTC 1140
QY 1153 TTTGACAGATGT 1164
Db 1141 TTTGACAGATGT 1152
```

```
RESULT 3
LOCUS CR623789
DEFINITION full-length cDNA clone CS01079YL05 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR623789
VERSION CR623789.1. GI:50504596
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1375)
AUTHORS Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
```

REMARK

Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1375)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

- Web : www.genoscope.cns.fr
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

FEATURES

source

1..1375
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01079YL05"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN

Query Match 82.7%; Score 996.8; DB 3; Length 1375;

Best Local Similarity 98.2%; Pred. No. 1.7e-269;

Matches 1127; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

```
QY 70 CTTCTCTCTGAGGAACCCGGCGATATGGCTGCGCTGTGCCCCCGCGCATTTCTCTC 129
Db 1 CTTCTCTCTGAGGAACCCGGCGATATGGCTGCGCTGTGCCCCCGCGCATTTCTCTC 60
QY 130 CCCGCTGTTT-CCCTTCTCTGGGTTCTCTCTCTCTCTCTCCGCTCCGCAATGGCGGAGCGG 188
Db 61 CCCGCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
QY 189 CTTGACACCAA-GGCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCAATCCCCA 247
Db 121 CTTGACACCAAAGGGGCGCCCTTCCCTGGATACGGTCACTTTCTACAGGTCAATCCCCA 180
QY 248 AAGCAAGTTCTCT-TGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAAGACAGATGA 306
Db 181 AAGCAAGTTCTCTTGGTGAAGTTTCGACACCCAGTACCCCTACGGTGAAGACAGATGA 240
QY 307 GTTCAAGCGTCTT-CTGAAAACTCGGTTTCAGCGATGATCTTTGGTGGCAGAGTGGG 365
Db 241 GTTCAAGCGTCTTGTGTAATACTCGGTTTCAGCGATGATCTTTGGTGGCAGAGTGGG 300
QY 366 GATCTCAGATTAT-GTGACAGCTGAACATGGAGCTGAGTGAGAAATACAGCTGGACAA 424
Db 301 GATCTCAGATTATGGTGAACATGGAGCTGAGTGAGAAATACAGCTGGACAA 360
QY 425 AGAGAGCTACCCA-TCTTCTACCTCTTCCGGATGGGACCTTTGAGAACCCAGTCCCCATA 483
Db 361 AGAGAGCTACCCAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420
QY 484 CACTGGGCGAGTT-AGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCGAAGGGGTCTACCT 542
Db 421 CACTGGGCGAGTTAAGGTTGGAGCCATCCAGCGCTGGCTGAAGGGGCGAAGGGGTCTACCT 480
QY 543 AGGTATGCTGG-TGCCTGCTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGGCTC 601
Db 481 AGGTATGCTGGTTCCTGTATACGACGCCCTTGGCCGGGGAGTTCATCAGGGGCTC 540
QY 602 TGGTGTGGAGG-CGCGCAGGCGCTTCTGAAGCAGGGGCAAGATACCTCTCAAGTGTGA 660
Db 541 TGGTGTGGAGGCCCGCCAGGCGCTTCTGAAGCAGGGGCAAGATACCTCTCAAGTGTGA 600
QY 661 GGAGACTCAGAAG-AGTGGGCGGAGCAATACCTCTGAAGATCATGGGGAAGATCTTTAGACCA 719
Db 601 GGAGACTCAGAAGAGTGGGCGGAGCAATACCTCTGAAGATCATGGGGAAGATCTTTAGACCA 660
QY 720 AGGGAGACATT-CCAGCATCAGAGATGACAGGATCCCGAGGCTGATTGAGAGAACAA 778
```

```
Db 661 AGGGAGGACTTCCAGCATCAGAGATGACACGATCGCCAGGCTGATTGAGAAGCA 720
Qy 779 GATGAGTCAC-GGCAGAGAGGAGCTCCAGAGAGCTTAAACATCTCGACTGCTTCCA 837
Db 721 GATGAGTCACGGAAGAAGAGGAGCTCCAGAGAGCTTAAACATCTCGACTGCTTCCA 780
Qy 838 GAAGAAGGGGGCCGAGAAAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTG 897
Db 781 GAAGAAGGGGGCCGAGAAAGAGGAGCTGTAAAAAGGCTGTCTGTGATTTTCCAGGGTTG 840
Qy 898 GTGGGGGTAGGAGGGGAGAGTTAACTGCTGGCTGTGTCANTCCCTTGTGGAAATATAAGG 957
Db 841 GTGGGGGTAGGAGGGGAGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGG 900
Qy 958 GGYMSKGGGAAAGWGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGA 1017
Db 901 GGTAGTGGGAAAGTGGTAAACCCAGATTTCTGAGCCCTGAGTATGCTGGACATTGA 960
Qy 1018 TGCTAAATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTAC 1077
Db 961 TGCTAAATGACCATGCTTGGGATGCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTAC 1020
Qy 1078 TCAGTGTGCTGTGAAATGACACCTCAGAGGAATGAGTGTCTATAGAGGAGAGAGG 1137
Db 1021 TCAGTGTGCTGTGAAATGACACCTCAGAGGAATGAGTGTCTATAGAGGAGAGAGG 1080
Qy 1138 TGTACTGCCAGGCTTTTGACAGATGTAATTTCTCAATTAAGTTTTCAGTGTTTGG 1197
Db 1081 TGTACTGCCAGGCTTTTGACAGATGTAATTTCTCAATTAAGTTTTCAGTGTTTGG 1140
Qy 1198 TTAAGTGG 1205
Db 1141 TTAAGTGG 1148

RESULT 4
LOCUS CR617833
DEFINITION full-length cDNA clone CS0DL010YA17 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617833.1 GI:50498640
VERSION HTC; CNSLT cDNA.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1139)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1139)
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
FEATURES
1..1139
source
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CS0DL010YA17"
/tissue_type="B cells (Ramos cell line) Cot 25-normalized"
```

```
ORIGIN
Query Match 82.0%; Score 987.8; DB 3; Length 1139;
Best Local Similarity 98.2%; Pred No. 5.7e-267;
Matches 1118; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

Qy 75 CTCTGAGGAACCCGGCGATATGGCTGCCGCTGTGCCCGCGCGCATTTCTCTCCCGC 134
Db 1 CTCTGAGGAACCCGGCGATATGGCTGCCGCTGTGCCCGCGCGCATTTCTCTCCCGC 60
Qy 135 TGCTT-CCCTTCTCTCTGGGCTTCTGCTCTCTCCGCTCCGCATGGCGGAGCGGCTGC 193
Db 61 TGCTTCCCTTCTCTGGGCTTCTGCTCTCTCCGCTCCGCATGGCGGAGCGGCTGC 120
Qy 194 ACACCAA-GGCGCCCTTCCCTCGATACGGTCACTTTCTCAAGGTCAATTCCTCCAAAAGCA 252
Db 121 ACACCAAAGGGCGCCCTTCCCTCGATACGGTCACTTTCTCAAGGTCAATTCCTCCAAAAGCA 180
Qy 253 AGTTTCGTC-TGGTGAAGTTCCGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCA 311
Db 181 AGTTTCGTTGGTGAAGTTCCGACACCCAGTACCCCTACGGTGAGAAGCAGGATGAGTTCA 240
Qy 312 AGCTCTTT-CTGAAAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCT 370
Db 241 AGCTCTTTGGTGA AAACTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCT 300
Qy 371 CAGATTAT-GTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGCACAAAAGCA 429
Db 301 CAGATTATGTTGACAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGCACAAAAGCA 360
Qy 430 GCTACCCA-TCTTCTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACATG 488
Db 361 GCTACCCAGTCTTCTACCTCTTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACATG 420
Qy 489 GGGCAGTT-AGTTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTA 547
Db 421 GGGCAGTTAAGGTTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACCTAGGTA 480
Qy 548 TGCTCTGG-TGCTCCCTGTATACGACGCCCTTGGCGGGAGTTTCATCAGGCGCTCTGGTG 606
Db 481 TGCTCTGGTGGCTGCTGTATACGACGCCCTTGGCGGGAGTTTCATCAGGCGCTCTGGTG 540
Qy 607 TGGAGG-CCGCCAGGCCCTTGTGAAGCAGGGCGCAAGATAAACCCTCAAGTGTGAAGGAGA 665
Db 541 TGGAGGCCGCCAGGCCCTTGTGAAGCAGGGCGCAAGATAAACCCTCAAGTGTGAAGGAGA 600
Qy 666 CTCAGAAG-AGTGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 724
Db 601 CTCAGAAGAGTGGGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGG 660
Qy 725 AGCACTT-CCAGCATCAGAGATGACGGATCGCCAGGCTGATTTGAGAAGAACAGATGA 783
Db 661 AGGACTTCCCAGCATCAGAGATGACGGATCGCCAGGCTGATTTGAGAAGAACAGATGA 720
Qy 784 GTGAC-GGCAGAAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCCCTCCAGAGA 842
Db 721 GTGACGGGAAGAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCTCCCTCCAGAGA 780
Qy 843 AGGGGGCGAGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGG 902
Db 781 AGGGGGCGAGAAAGAGGAGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTTGGTGG 840
Qy 903 GGTAGGAGGGGAGAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGYMS 962
Db 841 GGTAGGAGGGGAGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGGTAG 900
Qy 963 KGGGAAAAGWGGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGCATTTGATGCTA 1022
Db 901 TGGGAAAAGTGTACTAAACCCAGATTTCTGAGCCCTGAGTATGCTGCAGATTGATGCTA 960
Qy 1023 ACATGACCATGCTTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCACTTACTCAGG 1082
```

```
Db 961 ACATGACCATGCTTGGGATGTCCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGG 1020
Qy 1083 TGGCTGGTGAATGACACCTCAGAGGATGAGTCTATAGAGAGAGAGAGAGTGTAC 1142
Db 1021 TGGCTGGTGAATGACACCTCAGAGGATGAGTCTATAGAGAGAGAGAGAGTGTAC 1080
Qy 1143 TGGCCAGGTCTTTGACAGATGTAATTCATTCAATTAAGTTTCAGTGTGTTGTTAA 1201
Db 1081 TGGCCAGGTCTTTGACAGATGTAATTCATTCAATTAAGTTTCAGTGTGTTGTTAA 1139

RESULT 5
LOCUS CR602804 1119 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DC009YD12 of Neuroblastoma Cot
25-normalized of Homo sapiens (human).
ACCESSION CR602804
VERSION CR602804.1 GI:50483611
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1119)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1119)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and SclI sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
1..1119
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC009YD12"
/tissue_type="Neuroblastoma Cot 25-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 80.3%; Score 967.8; DB 3; Length 1119;
Best Local Similarity 98.1%; Pred. No. 2.5e-261;
Matches 1098; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
Qy 88 CGGCGATATGCTGCGCTGTGCGCGCGCGCGCATTTCTCCCGCTGCTT-CCCTTCT 146
Db 1 CGGCGATATGCTGCGCTGTGCGCGCGCGCATTTCTCCCGCTGCTTCCCTTCT 60

Qy 147 CTGGGCTTCTGCTCTCTCGCTCCGATGGCGGCGGCGCTGCACACAA-GGCGC 205
Db 61 CTGGGCTTCTGCTCTCTCGCTCCGATGGCGGCGGCGCTGCACACAAAGGGCGC 120
Qy 206 CTTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTGCTG- 264
Db 121 CTTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTGCTG 180
Qy 265 GAAGTTGACACCCAGTACCCCTACGGTGAGAGAGAGAGATGTTCAAGCGTCTT-CTGA 323
Db 181 GAAGTTGACACCCAGTACCCCTACGGTGAGAGAGAGATGTTCAAGCGTCTTGTCTGA 240
Qy 324 AAATCGGCTTCCAGGATGATCTCTTGGTGACAGGTGGGATCTCAGATTAT-CTGA 382
```

```
241 AAATCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTATGTTGA 300
Qy 383 CAAGCTGAACATGAGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGAGCTACCCCA-TCCT 441
Db 301 CAAGCTGAACATGAGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGAGCTACCCAGTCTT 360
Qy 442 CTACCTCTTCCGGATGGGAGCTTTTGAAACCCAGTCCCATACACCTGGGGCAGTT-AGT 500
Db 361 CTACCTCTTCCGGATGGGAGCTTTTGAGAACCCAGTCCCATACACCTGGGGCAGTTAAGGT 420
Qy 501 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTTACCTAGGTATGCCTGG-TGCCT 559
Db 421 TGGAGCCATCCAGCGCTGGCTGAAGGGGCAAGGGGCTTACCTAGGTATGCCTGGTTCCT 480
Qy 560 GCCTGTATACAGAGCTTGGCGGGGAGTTTCAATCAGGGCTCTGTGTGGAGG-CCGCA 618
Db 481 GCCTGTATACAGAGCTTGGCGGGGAGTTTCAATCAGGGCTCTGTGTGGAGGCGCCGCA 540
Qy 619 GGCCCTCTTGAAGCAGGGGCAAGATAAATCTCAAGTGTGAAGGAGACTCAGAAG-AGTG 677
Db 541 GGCCCTCTTGAAGCAGGGGCAAGATAAATCTCAAGTGTGAAGGAGACTCAGAAGAGTG 600
Qy 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 736
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGC 660
Qy 737 ATCAGAGATGACAGGATCGCCAGCTGATTTGAGAGAAACAAGATGAGTGAC-GCAGAA 795
Db 661 ATCAGAGATGACAGGATCGCCAGCTGATTTGAGAGAAACAAGATGAGTGACGCGGAAGAA 720
Qy 796 GGAGGAGCTCCAGAGAGAGCTTAAACATCTCACTGCCTTCCAGAGAGAGGGGGCCGAGAA 855
Db 721 GGAGGAGCTCCAGAGAGAGCTTAAACATCTCACTGCCTTCCAGAGAGAGGGGGCCGAGAA 780
Qy 856 AGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGA 915
Db 781 AGAGGAGCTGTAAAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGA 840
Qy 916 NAGTTAACTCTCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGYMSKGGAAAGWGGT 975
Db 841 GAGTTAACTCTCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGTAGTGGGAAAGTGGT 900
Qy 976 ACTAACCCAGCACTTCTGAGCCCTGAGTATGCTGGACATTCATGCTAACATGACCATGCT 1035
Db 901 ACTAACCCAGCACTTCTGAGCCCTGAGTATGCTGGACATTCATGCTAACATGACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGTCTGGGATAGCTTGGAGCACTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGTCTGGGATAGCTTGGAGCACTTACTCAGGTGGCTGGTGAAT 1020
Qy 1096 GACACCTCAGAGGAATGAGTCTATAGAGAGAGAGAGAGTGTACTGCCCAGGTCTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTCTATAGAGAGAGAGAGAGTGTACTGCCCAGGTCTTT 1080
Qy 1156 GACAGATGATTTCTCAATTCATTAAGTTTTCAGTGT 1194
Db 1081 GACAGATGATTTCTCAATTCATTAAGTTTTCAGTGT 1119

RESULT 6
LOCUS CR611287 1119 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DI072YH08 of Placenta Cot 25-normalized
of Homo sapiens (human).
ACCESSION CR611287
VERSION CR611287.1 GI:50492094
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1119)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
AUTHORS
```

```
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished
REMARK     Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue
REFERENCE  2 (bases 1 to 1119)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT    1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES   Location/Qualifiers
            source
              1..1119
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODI072YH08"
                /tissue_type="Placenta Cot 25-normalized"
                /plasmid="pCMVSPORT_6"
ORIGIN
Query Match      80.3%; Score 967.8; DB 3; Length 1119;
Best Local Similarity 98.1%; Pred. No. 2.5e-261;
Matches 1098; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
QY 70 C T C C C C T C T G C A G G A A C C G G C G A T A T G G C T G C C G C T G T G C C G C G C G C G C A T T T C T C T C 129
DB 1 C C T C C C C T C T G C A G G A A C C G G C G A T A T G G C T G C C G C T G T G C C C G C G C G C A T T T C T C T C 60
QY 130 C C C G C T G C T T - C C C T T C T C T G G G T T C C T G C T C T C C G C T C C G C A T G C G G C A G G G 188
DB 61 C C C G C T G C T T C C C C T T C C T G G G T T C C T G C T C T C C G C T C C G C A T G C G G C A G G G 120
QY 189 C T G C A C A C C A A - G G C G C C T T C C C C T G A T A C G G T C A C T T T C T A C A A G G T C A T T C C C A A 247
DB 121 C C T G C A C A C C A A G G G G C C C T T C C C C T G A T A C G G T C A C T T T C T A C A A G G T C A T T C C C A A 180
QY 248 A A G C A A G T T C G T C - T G S T G A A G T T C G A C A C C C A G T A C C C C T A C G G T G A G A A C G C A G G A T G A 306
DB 181 A A G C A A G T T C G T T G T G T G A G T T C G A C A C C C A G T A C C C T A C G G T G A G A A C G C A G G A T G A 240
QY 307 G T T C A A G C G T C T T - C T G A A A A C T C G G C T T C A G C A T G A T C T T T G G T G G C A G A G G T G G G 365
DB 241 G T T C A A G C G T C T T G C T G A A A A C T C G G C T T C A G C A T G A T C T T T G T G G C A G A G G T G G G 300
QY 366 G A T C T C A G A T A T - C T G C A A G C T C A A C A T G G A G C T G A G T G A G A A T A C A A G C T G G A C A A 424
DB 301 G A T C T C A G A T A T G T G T G A C A G C T G A A C A T G G A G C T G A G T G A G A A T A C A A G C T G G A C A A 360
QY 425 A G A G A G C T A C C C A - T C T T T A C T C T T C G G G A T G G G A C T T T G A G A A C C C A G T C C C A T A 483
DB 361 A G A G A G C T A C C C A G C T T T C A C T C T T C G G G A T G G G A C T T T G A G A A C C C A G T C C C A T A 420
QY 484 C A C T G G G C A G T T - A G G T T G A G C C A T C A G C G C T G G C T G A A G G G G C A A G G G T C T A C C T 542
DB 421 C A C T G G G C A G T T A A G G T T G G A G C C A T C A G C G C T G G C T G A A G G G G C A A G G G T C T A C C T 480
QY 543 A G G T A T G C T G - T G C C T G C C T G T A T A C G A C C C T G G C G G G A G T T C A T C A G G G C T C 601
DB 481 A G G T A T G C C T G G T G C C T G T A T A C A C C C C T G G C G G G A G T T C A T C A G G G C T C 540
QY 602 T G G T G T G G A G G - C C G C C A G G C C C T T T G A A G C A G G G C A A G A T A C C T C T C A A G T G T G A A 660
DB 541 T G G T G T G A G G C C C G C C A G G C C C T C T T G A A G C A G G G C A A G A T A C C T C T C A A G T G T G A A 600
QY 661 G G A G A C T C A G A A G - A G T G G G C C G A G C A T A C C T G A A G A T C A T G G G G A G A T C T T A G A C C A 719
DB 601 G G A G A C T C A G A A G A G T G G G C C G A G C A A T A C C T G A A G A T C A T G G G G A A G A T C T T A G A C C A 660
```

```
QY 720 A G G G G A G C A C T T - C C A G C A T C A G A G A T G A C A C G G A T C C C A G G C T G A T T G A G A A G A C A A 778
DB 661 A G G G G A G C A C T T C C C A C A T C A G A G A T G A C A C G G A T C C C A G G C T G A T T G A G A A G A C A A 720
QY 779 G A T G A G T G A C - G G C A G A A G A G A G A G C T C C A G A G A G C T T A A A C A T C C T G A C T G C C T T C C A 837
DB 721 G A T G A G T G A C G G G A A G A G A G A G C T C C A G A G A G C T T A A A C A T C C T G A C T G C C T T C C A 780
QY 838 G A G A A G G G G C C G A G A A A G A G A G C T G T A A A A G G C T G T C T G T A T T T T C C A G G G T T T G 897
DB 781 G A G A A G G G G C C C A G A A A G A G A G A G C T G T A A A A G G C T G T C T G T A T T T T C C A G G G T T T G 840
QY 898 G T G G G G T A G G A G G G G A N A G T T A A C C T G C T G G C T G T G A N T C C C T T G T G A A T A T A A G G G 957
DB 841 G T G G G G T A G G A G G G G A G A G T T A A C C T G C T G G C T G T G A G T C C C T T G T G A A T A T A A G G G 900
QY 958 G G Y M S K G G A A A A W G G T A C T A A C C A C G A T T C T G A G C C C T G A G T A T G C C T G G A C A T T G A 1017
DB 901 G G T A G T G G G A A A G T G G T A C T A A C C A C G A T T C T G A G C C C T G A T G A T G C C T G G A C A T T G A 960
QY 1018 T G C T A A C A T G A C C A T G C T T G G G A T G T C T A G C T G G T C T G G G A T A G C T G G A G C A C T T A C 1077
DB 961 T G C T A A C A T G A C C A T G C T T G G G A T G T C T A G C T G G T C T G G G A T A G C T G G A G C A C T T A C 1020
QY 1078 T C A G G T G C T G G T G A A A T G A C A C C T C A G A A G A A T G A G T G C T A T A G A G A G A G A G A G G A G 1137
DB 1021 T C A G G T G C T G G T G A A A T G A C A C C T C A G A A G A A T G A G T G C T A T A G A G A G A G A G A G A G 1080
QY 1138 T G T A C T G C C C A G G T C T T T G A C A G A T G A A T T C T C A T T C A 1176
DB 1081 T G T A C T G C C C A G G T C T T T G A C A G A T G A A T T C T C A T T C A 1119
RESULT 7
LOCUS     CR604444      1111 bp      mRNA      linear      HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODI043YB01 of Placenta Cot 25-normalized
            of Homo sapiens (human) .
ACCESSION CR604444
VERSION   CR604444.1 GI:50485251
KEYWORDS  HTC; CNSLUT_cDNA.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1111)
AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE     Full-length cDNA libraries and normalization
JOURNAL   Unpublished
REMARK    Contact : Feng Liang Email : fliang@lifetech.com URL :
           http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
           Faraday Avenue
           2 (bases 1 to 1111)
REFERENCE 2 (bases 1 to 1111)
AUTHORS    Genoscope.
TITLE      Direct Submission
JOURNAL    Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
           BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr
           - Web : www.genoscope.cns.fr)
COMMENT    1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
           end enriched, double-strand cDNA was digested with Not I and cloned
           into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
           was normalized. Library was constructed by Life Technologies, a
           division of Invitrogen.
FEATURES   Location/Qualifiers
            source
              1..1111
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CSODI043YB01"
                /tissue_type="Placenta Cot 25-normalized"
                /plasmid="pCMVSPORT_6"
ORIGIN
```

Query Match 79.7%; Score 959.8; DB 3; Length 1111;
Best Local Similarity 98.1%; Pred. No. 4.5e-259;
Matches 1090; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

QY 88 CGCGGATATGGCTGGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCTTTCT 146
DB 1 CGCGGATATGGCTGGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTTCT 60

QY 147 CTGGGCTTCCTGCTCTCTCCGCTCCGATGGCGGCGGCTGCACACAA-GGCGC 205
DB 61 CTGGGCTTCCTGCTCTCTCCGCTCCGATGGCGGCGGCTGCACACAAAGGGCGC 120

QY 206 CTTTCCCTGGATAGCTGCTTCTTACAAAGTCTTCCAAAGCAAGTTCGTC-TGGT 264
DB 121 CTTTCCCTGGATAGCTGCTTCTTACAAAGTCTTCCAAAGCAAGTTCGTC-TGGT 180

QY 265 GAAGTTCCGACACCCAGTACCTTCTCCGCTCCGATGGCGGCGGCTGCACACAA-GGCGC 323
DB 181 GAAGTTCCGACACCCAGTACCTTCTCCGCTCCGATGGCGGCGGCTGCACACAAAGGGCGC 240

QY 324 AAATCGGCTTCAGCGATGATCTTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
DB 241 AAATCGGCTTCAGCGATGATCTTCTTGGTGGCAGAGTGGGATCTCAGATTATGTGA 300

QY 383 CAAGCTGACATGGAGCTGAGTGAGAAATACAGCTGGAAGAGAGCTACCCA-TCTT 441
DB 301 CAAGCTGACATGGAGCTGAGTGAGAAATACAGCTGGAAGAGAGCTACCCAAGTCTT 360

QY 442 CTACCTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACCTGGGCGAGTT-AGT 500
DB 361 CTACCTCTTCCGGATGGGACTTTGAGAACCCAGTCCCATACCTGGGCGAGTTAAGT 420

QY 501 TGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGCTTACCTAGTATGCTTGG-TGCT 559
DB 421 TGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGCTTACCTAGTATGCTTGGTGGCT 480

QY 560 GCCTGTATACGAGCGCTGGCGGGAGTTTATCAGGGCTCTGTGTGGAGG-CCGCA 618
DB 481 GCCTGTATACGAGCGCTGGCGGGAGTTTATCAGGGCTCTGTGTGGAGGCGCCGCA 540

QY 619 GGCCTCTTGAAGAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGCTCAGAAG-AGTG 677
DB 541 GGCCTCTTGAAGAGGGGCAAGATAAAGCTCTCAAGTGTGAAGGAGCTCAGAAGAGTG 600

QY 678 GCGGAGCAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGC 736
DB 601 GCGGAGCAATACCTGAAGATCATGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGC 660

QY 737 ATCAGAGATGACCGGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAC-GGCAGAA 795
DB 661 ATCAGAGATGACCGGATGCGCAGGCTGATTTGAGAGAAACAGATGAGTGAGGAGAA 720

QY 796 GGAGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCGCAGAA 855
DB 721 GGAGAGCTCCAGAGAGCTTAAACATCTGACTGCTTCCAGAGAAAGGGGCGCAGAA 780

QY 856 AGAGAGCTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 915
DB 781 AGAGAGCTGTAAGAGGCTGTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGGAGGGGA 840

QY 916 NAGTTAACTGTGCTGTGANTCCCTTGTGGAATATTAAGGGGYSKGGGAAAGWGGT 975
DB 841 GAGTTAACTGTGCTGTGANTCCCTTGTGGAATATTAAGGGGYSKGGGAAAGWGGT 900

QY 976 ACTACCCAGATCTGAGCCCTGAGTATGCTTGGACATTTGATGCTTACATCACCATGCT 1035
DB 901 ACTACCCAGATCTGAGCCCTGAGTATGCTTGGACATTTGATGCTTACATCACCATGCT 960

QY 1036 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1095
DB 961 TGGGATGCTCTAGCTGGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGGTGAAT 1020

QY 1096 GACACCTCAGAGGAAGTGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGCTCTTT 1155
DB 1021 GACACCTCAGAGGAAGTGTCTATAGAGAGGAGAGGAGTGTACTGCCAGGCTCTTT 1080

QY 1156 GACAGATTAATTTCTCAATTAAGTCTT 1186
DB 1081 GACAGATTAATTTCTCAATTAAGTCTT 1111

RESULT 8
LOCUS CR617687 1111 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CSODL001YK02 of B cells (Ramos cell line)
Cot 25-normalized of Homo sapiens (human).
ACCESSION CR617687.1 GI:50498494
VERSION CR617687.1
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1111)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1111)
REFERENCE 2 (bases 1 to 1111)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1111
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL001YK02"
/tissue type="B cells (Ramos cell line) Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 79.7%; Score 959.8; DB 3; Length 1111;
Best Local Similarity 98.1%; Pred. No. 4.5e-259;
Matches 1090; Conservative 5; Mismatches 4; Indels 12; Gaps 12;

QY 88 CGCGGATATGGCTGGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCTTTCT 146
DB 1 CGCGGATATGGCTGGCGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTTCT 60

QY 147 CTGGGCTTCCTGCTCTCTCCGCTCCGATGGCGGCGGCTGCACACAA-GGCGC 205
DB 61 CTGGGCTTCCTGCTCTCTCCGCTCCGATGGCGGCGGCTGCACACAAAGGGCGC 120

QY 206 CTTTCCCTGGATAGCTGCTTCTTACAAAGTCTTCCAAAGCAAGTTCGTC-TGGT 264
DB 121 CTTTCCCTGGATAGCTGCTTCTTACAAAGTCTTCCAAAGCAAGTTCGTC-TGGT 180

QY 265 GAAGTTCCGACACCCAGTACCTTCTCCGCTCCGATGGCGGCGGCTGCACACAA-GGCGC 323
DB 181 GAAGTTCCGACACCCAGTACCTTCTCCGCTCCGATGGCGGCGGCTGCACACAAAGGGCGC 240

QY 324 AAATCGGCTTCAGCGATGATCTTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382
DB 241 AAATCGGCTTCAGCGATGATCTTCTTGGTGGCAGAGTGGGATCTCAGATTATGTGA 300


```
Qy 383 CAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGACTACCCA-TCTT 441
Db 301 CAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGACTACCCA-TCTT 360
Qy 442 CTACCTCTTCGGGATGGGGACTTTGAGAACCCAGTCCCATACATACACTGGGGCAGTT-AGGT 500
Db 361 CTACCTCTTCGGGATGGGGACTTTGAGAACCCAGTCCCATACATACACTGGGGCAGTTAAGGT 420
Qy 501 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTCTGG-TGCCT 559
Db 421 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTCTGG-TGCCT 480
Qy 560 GCCTGTATACGACGCTCGCTGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG-CCGCCA 618
Db 481 GCCTGTATACGACGCTCGCTGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGGCGGCCA 540
Qy 619 GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAG-AGTG 677
Db 541 GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAGAGTG 600
Qy 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGAGACTTCCACG 736
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGAGACTTCCACG 660
Qy 737 ATCAGAGATGACAGGATCGCCAGCTGATTTGAGAAGAACAGATGATGAC-GGCAGAA 795
Db 661 ATCAGAGATGACAGGATCGCCAGCTGATTTGAGAAGAACAGATGATGACGGGAGAA 720
Qy 796 GGAGAGCTCAGAAAGACTTAAACATCTGACTGCTTCCAGAAAGAGGGGGCGAGAA 855
Db 721 GGAGAGCTCAGAAAGACTTAAACATCTGACTGCTTCCAGAAAGAGGGGGCGAGAA 780
Qy 856 AGAGAGCTGTAAAGAGCTGTCTGATTTTCAGAGGTTTGGTGGGGTAGGGAGGGGA 915
Db 781 AGAGAGCTGTAAAGAGCTGTCTGATTTTCAGAGGTTTGGTGGGGTAGGGAGGGGA 840
Qy 916 NAGTTAACTGTGCTGTGANTCCCTTGTGGAATATAAGGGGGYMSKGGAAAGMGST 975
Db 841 GAGTTAACTGTGCTGTGAGTCCCTTGTGGAATATAAGGGGGTAGTGGAAAGTGST 900
Qy 976 ACTAACCCAGATTCAGACCTCAGTATGCTTGACATATGATGCTAATACCATGCT 1035
Db 901 ACTAACCCAGATTCAGACCTCAGTATGCTTGACATATGATGCTAATACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGTCTGGGATAGCTGGACACTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGTCTGGGATAGCTGGACACTTACTCAGGTGGCTGGTGAAT 1020
Qy 1096 GACACCTCAGAGGAATGAGTGTCTATAGAGGAGAGAGTGTACTGCCAGGTCTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTGTCTATAGAGGAGAGAGTGTACTGCCAGGTCTTT 1080
Qy 1156 GACAGATGTAATCTCATTTAAAGTTT 1186
Db 1081 GACAGATGTAATCTCATTTAAAGTTT 1111
```

RESULT 9
CR603906
LOCUS
DEFINITION
full-length cDNA clone CS0DH002YC21 of T cells (Jurkat cell line)
of Homo sapiens (human).
ACCESSION
CR603906
VERSION
CR603906.1 GI:50484713
KEYWORDS
HTC; CNSLT_cDNA.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1110)
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished

```
REMARK  
Contact : Feng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue  
2 (bases 1 to 1110)  
Genoscope.  
Direct Submission  
TITLE  
JOURNAL  
COMMENT  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
FEATURES  
Location/Qualifiers  
1..1110  
source  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="CS0DH002YC21"  
/tissue_type="T cells (Jurkat cell line)"  
/plasmid="pCMVSPORT_6"  
ORIGIN  
Query Match 79.6%; Score 958.8; DB 3; Length 1110;  
Best Local Similarity 98.1%; Pred. No. 8.6e-259;  
Matches 1089; Conservative 5; Mismatches 4; Indels 12; Gaps 12;  
Qy 88 CGGCGATATGGTCCGCTGTGCCCGCGCGCATTTCTCCCGCGTCTT-CCCTTCT 146  
Db 1 CGGCGATATGGTCCGCTGTGCCCGCGCGCATTTCTCTCCCGCTTCTCCCTTCT 60  
Qy 147 CCTGGCTTCTCTCTCTCCGCTCCGCTGCGCGAGCGGCTGCACACCAA-GGCGC 205  
Db 61 CTTGGCTTCTCTCTCTCTCCGCTCCGCTGCGCGAGCGGCTGCACACCAAAGGGCGC 120  
Qy 206 CTTTCCCTGGATACGCTCACTTTCTCAAGGTCAATTTCCCAAAGCAAGTTCTGTC- 264  
Db 121 CTTTCCCTGGATACGCTCACTTTCTCAAGGTCAATTTCCCAAAGCAAGTTCTGTC-TG 180  
Qy 265 GAAGTTGCACACCCAGTACCCCTACGCTGAGAGAGAGAGTTCAGCGCTCTT-CTGA 323  
Db 181 GAAGTTGCACACCCAGTACCCCTACGCTGAGAGAGAGAGTTCAGCGCTCTTCTGA 240  
Qy 324 AAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTAT-GTGA 382  
Db 241 AAATCTGGCTTCCAGCGATGATCTCTTGGTGGCAGAGTGGGATCTCAGATTATGTGA 300  
Qy 383 CAAGCTGAACATGAGCTGAGTGAATAACAAGCTGGAACAAGAGAGCTACCCA-TCTT 441  
Db 301 CAAGCTGAACATGAGCTGAGTGAATAACAAGCTGGAACAAGAGAGCTACCCAAGTCTT 360  
Qy 442 CTACCTCTTCGGGATGGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTT-AGGT 500  
Db 361 CTACCTCTTCGGGATGGGGACTTTGAGAACCCAGTCCCATACACTGGGGCAGTTAAGGT 420  
Qy 501 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTCTGG-TGCCT 559  
Db 421 TGGAGCCATCAGCGCTCGCTGAAGGGGCAAGGGTCTACCTAGTATGCTCTGGTTGCT 480  
Qy 560 GCCTGTATACGACGCTCGCTGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGG-CCGCCA 618  
Db 481 GCCTGTATACGACGCTCGCTGGCGGGAGTTTCATCAGGGCTCTGTGTGGAGGCGGCCA 540  
Qy 619 GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAG-AGTG 677  
Db 541 GGCCCTCTTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGAGACTCAGAAGAGTG 600  
Qy 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGAGACTT-CCAGC 736  
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGAGACTTCCACG 660  
Qy 737 ATCAGAGATGACAGGATCGCCAGCTGATTTGAGAAGAACAGATGATGAC-GGCAGAA 795
```



```
Db 661 ATCAGAGATGACAGGATCGCCAGGCTGATTTGAGAGAACACAGATGATGACGGAGAA 720
Qy 796 GGAGGAGCTCCAGAGAGCTTAAACATCTGATCCCTTCCAGAGAAAGGGGCGAGAA 855
Db 721 GGAGGAGCTCCAGAGAGCTTAAACATCTGATCGCTTCCAGAGAAAGGGGCGAGAA 780
Qy 856 AGAGGAGCTGTAAGAGGCTGCTCTGATTTTCCAGGGGTTTGGTGGGGTAGGGAGGGA 915
Db 781 AGAGGAGCTGTAAGAGGCTGCTCTGATTTTCCAGGGGTTTGGTGGGGTAGGGAGGGA 840
Qy 916 NAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGYNMKGSGGAAAGWGT 975
Db 841 GAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGTTAGTGGGAAAGTGT 900
Qy 976 ACTAACCCAGCATTCGTAGCCCTGATGATGCTCGACATTTGATGCTAAATGACCATGCT 1035
Db 901 ACTAACCCAGCATTCGTAGCCCTGATGATGCTCGACATTTGATGCTAAATGACCATGCT 960
Qy 1036 TGGGATGCTCTAGCTGTCTGGGATAGCTGGGACCTTACTCAGGTGGCTGGTGAAT 1095
Db 961 TGGGATGCTCTAGCTGTCTGGGATAGCTGGGACCTTACTCAGGTGGCTGGTGAAT 1020
Qy 1096 GACACCTCAGAGGAATGAGTCTATAGAGGAGAGGAGTGTACTGCCAGGTCTTT 1155
Db 1021 GACACCTCAGAGGAATGAGTCTATAGAGGAGAGGAGTGTACTGCCAGGTCTTT 1080
Qy 1156 GACAGATGTAATTCATTCATTAAGTT 1185
Db 1081 GACAGATGTAATTCATTCATTAAGTT 1110
```

```
RESULT 10
BX403042
LOCUS BX403042 Homo sapiens B CELLS (RAMOS CELL LINE) linear EST 28-APR-2004
DEFINITION Homo sapiens cDNA clone CSODL010YA17 5-PRIME, mRNA sequence.
ACCESSION BX403042
VERSION BX403042.2 GI:46844967
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1137)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30611153.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6716.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna/cna=CSIAL003ZC09QPI&c=6716.f.
Location/Qualifiers
1. 1137
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODL010YA17"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/cell_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
```

FEATURES

source

```
ORIGIN
Query Match 77.0%; Score 928.4; DB 5; Length 1137;
Best Local Similarity 94.8%; Pred. No. 3.3e-250;
Matches 1080; Conservative 23; Mismatches 22; Indels 14; Gaps 14;
Qy 75 CTCTGCAGGAACCCGGCGGATATGGCTGCCGTGTGCCCCCGCGGCATTTCTCTCCCCGC 134
Db 1 CTCTGCAGGAACCCGGCGGATATGGCTGCCGTGTGCCCCCGCGGCATTTCTCTCCCCGC 60
Qy 135 TGCCTT-CCCTTCTCTCTGGGCTTCTGCTCTCTCGCTCCGCTCCGCTCCGCGGCGGCGGCTGC 193
Db 61 TGCCTTCCCTTCTCTCTGGGCTTCTGCTCTCTCGCTCCGCTCCGCTCCGCGGCGGCGGCTGC 120
Qy 194 ACACCAA-GGCGCCCTTCCCTTGGATACGGTCACTTTCTACAAGGTCAATTCCTCAAAAGCA 252
Db 121 ACACCAAGGGCGCCCTTCCCTTGGATACGGTCACTTTCTACAAGGTCAATTCCTCAAAAGCA 180
Qy 253 AGTTCGTC-TGGTGAAGTTTCACACCCAGTACCCTACGGTGAGAGCAGGATGATTTCA 311
Db 181 AGTTCGTC-TGGTGAAGTTTCACACCCAGTACCCTACGGTGAGAGCAGGATGATTTCA 240
Qy 312 AGCTCTTT-CTGAAAACCTGGCTTCCAGCGATGATCTCTTGGTGCAGAGTGGGGATCT 370
Db 241 AGCTCTTTCTGTAAGAACCTGGCTTCCAGCGATGATCTCTTGGTGCAGAGTGGGGATCT 300
Qy 371 CAGATTAT-GTGACAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGA 429
Db 301 CAGATTATGGTGACACAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGAACAAGAGA 360
Qy 430 GCTACCCA-TCTTCTACCTTCTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTG 488
Db 361 GCTACCCAGCTTCTTCTACCTTCTCCGGGATGGGACTTTTGAGAACCCAGTCCCATACACTG 420
Qy 489 GGGCAGTT-AGTTGGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGTCTACCTAGTA 547
Db 421 GGGCAGTTAAGTTTGGAGCCATCCAGCGCTGGCTGGAAGGGGCAAGGGTCTACCTAGTA 480
Qy 548 TGCCTGG-TGCCTGCTGTATACGACGCCCTGGCGGGGAGTTTCAATCAGGGCTCTGTGTG 606
Db 481 TGCCTGGTTGCTGCGGGTATACGACGCCCTGGCGGGGAGTTTCAATCAGGGCTCTGTGTG 540
Qy 607 TGGAGG-CCGCGCAGCCCTCTTTGAAGCAGGGGCAAGATAAACCCTCTCAAGTGTGAAGAGA 665
Db 541 TGGAGGCCCGCCAGGCCCTCTTTGAAGCAGGGGCAAGATAAACCCTCTCAAGTGTGAAGAGA 600
Qy 666 CTCAGAG-AGTGGCGCGAGCAATACCTGGAAGATCATGGGGAAGATCTTAGACCAAGGGG 724
Db 601 CTCAGAGAAGTGGGCGGAGCAATACCTGGAAGATCATGGGGAAGATCTTAGACCAAGGGG 660
Qy 725 AGCACTT-CCAGCATCAGAGATGACAGCGATCGCCAGGCTGATTGAGAGAACCAAGATGA 783
Db 661 AGGACTTCCAGCATCAGAGATGACAGCGATCGCCAGGCTGATTGAGAGAACCAAGATGA 720
Qy 784 GTGAC-GGCGAGAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCCCTTCCAGAGA 842
Db 721 GTGACGGGAAGAGGAGGAGCTCCAGAGAGCTTAAACATCTGACTGCCCTTCCAGAGA 780
Qy 843 AGGGGGCGAGAAAGAGAGGCTGTAAAGGCTGTCTGTGATTTTCCAGGGTTGGTGGG 902
Db 781 AGGGGGCGAGAAAGAGAGGCTGTAAAGGCTGTGTGTTATTTTCCAGGGTTGGTGGG 840
Qy 903 GGTAGGGGGGAGAGTTAACTCTGCTGTGANTCCCTTGTGGAATATAAGGGGGYMS 962
Db 841 GGTAGGGGGGAGAGTTAACTCTGCTGTGTTAGTTCTTGTGGAATATAAGGGGGTAG 900
Qy 963 KGGGAAAAGWGTACTAAACCCAGATTCAGAGCCCTCAGATGCTGCTGACATTTGATGCTA 1022
Db 901 TGGGAAAAGTGTACTAAACCCAGATTCAGAGCCCTCAGATGCTGCTGACATTTGATGCTA 960
Qy 1023 ACATGACCATCTCTGGGAGTGTCTAGCTGTGTGGGATAGCTGGAGCACTTACTCAGG 1082
```

digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized."

Db 961 AMATGACCATGCTGGGATGCTCTAGCTGCTGGGATAGCTGGAGCMCTTACTCAGG 1020
Qy 1083 TGGCTGGTGAATGACACCTCAGAGGAATGAGTCTATAGAGGAGGAGGAGTGTAC 1142
Db 1021 TGGCTGKTGAATGACAMCTCAGAGGAATGARTTCTAWA-AAAAGRRAAAAGAGTTTC 1079
Qy 1143 TGCCAGGTCTTTGACAGATGAATTCATTCATTAAATTAAGTTTCAGTGTTCGTTAA 1201
Db 1080 TGCCACGTCTTTTACAGKGTWATTTTWT-MATTAARATTCAGTKTTTGKTTAA 1137

RESULT 11
CR602432
LOCUS full-length cDNA clone CS0DF036YH02 of Petal brain of Homo sapiens (human) 1062 bp mRNA linear HTC 21-JUL-2004
DEFINITION
ACCESSION CR602432
VERSION CR602432.1 GI:50483239
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/InvitroGen Corporation 1600 Faraday Avenue
2 (bases 1 to 1062)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen.
FEATURES
Location/Qualifiers
1..1062
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF036YH02"
/tissue type="Fetal brain"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 75.6%; Score 910.8; DB 3; Length 1062;
Best Local Similarity 98.0%; Pred. No. 3e-245;
Matches 1041; Conservative 5; Mismatches 4; Indels 12; Gaps 12;
Qy 88 CGGCGATATGGTCGGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CCCTTCT 146
Db 1 CGGCGATATGGTCGGCTGTGCCCGCGCGCATTTCTCTCCCGCTGCTTCCCTTCT 60
Qy 147 CTTGGGCTTCTGCTCTCTCCGCTCCGCGCGCGCGCTGCGCACACAA-GGCGC 205
Db 61 CTTGGGCTTCTGCTCTCTCCGCTCCGCGCGCGCGCTGCGCACACAAAGGGCGC 120
Qy 206 CTTTCCCTGGATAGGTCACATTTCTACAGGTCATTCCTCCAAAGCAAGTTGCTC 264
Db 121 CTTTCCCTGGATAGGTCACATTTCTACAGGTCATTCCTCCAAAGCAAGTTGCTC 180
Qy 265 GAAGTTCACACCCAGTACCCCTAGGTCAGAGGAGGATGATTCAGCGCTT-CTGA 323
Db 181 GAAGTTCACACCCAGTACCCCTAGGTCAGAGGAGGATGATTCAGCGCTTCTGCTGA 240
Qy 324 AAATCTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGGATCTCAGATTAT-GTGA 382

Db 241 AAATCTCGGCTTCCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATTATGTTGA 300
Qy 383 CAAGCTGAAATGATGAGCTGAGTGAATATCAAGCTGGAACAAGAGAGCTACCCA-TCTTT 441
Db 301 CAAGCTGAAATGATGAGCTGAGTGAATATCAAGCTGGAACAAGAGAGCTACCCAAGTCTT 360
Qy 442 CTACTCTCTCCGGATGGGACCTTTGAGAAACCCAGTCCCATACACTGGGCGAGTT-AGGT 500
Db 361 CTACTCTCTCCGGATGGGACCTTTGAGAAACCCAGTCCCATACACTGGGCGAGTTAGGT 420
Qy 501 TGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGTCTTACTAGGTATGCTCTGG-TGCCT 559
Db 421 TGGAGCCATCAGCGCTGGCTGAAGGGCAAGGGTCTTACTAGGTATGCTCTGGTGTGCT 480
Qy 560 GCCTGTATAGACGCCCTGCCGGGAGTTTCATCAGGGCTCTCTGGTGTGGAGG-CCGCCA 618
Db 481 GCCTGTATAGACGCCCTGCCGGGAGTTTCATCAGGGCTCTCTGGTGTGGAGGCCCCGCA 540
Qy 619 GGCCTCTCTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAAG-AGTG 677
Db 541 GGCCTCTCTGAAGCAGGGGCAAGATAACCTCTCAAGTGTGAAGGAGACTCAGAAAGAGTG 600
Qy 678 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CAAGC 736
Db 601 GGCCGAGCAATACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCCGAGC 660
Qy 737 ATCAGAGATCAGCGGATCGCCAGGCTGATTGAGAACAACAAGATGATGAC-GGCAGAA 795
Db 661 ATCAGAGATCAGCGGATCGCCAGGCTGATTGAGAACAACAAGATGATGACGGGAAGAA 720
Qy 796 GGAGGAGCTCCAGAAAGAGCTTAAACATCTGACCTCCAGAAAGAGGGGGCCGAGAA 855
Db 721 GGAGGAGCTCCAGAAAGAGCTTAAACATCTGACCTCCAGAAAGAGGGGGCCGAGAA 780
Qy 856 AGAGGAGCTGTAAAGAGCTGTCTGTGATTTTCAGGGTTTGGTGGGGTATAGGAGGGGA 915
Db 781 AGAGGAGCTGTAAAGAGCTGTCTGTGATTTTCAGGGTTTGGTGGGGTATAGGAGGGGA 840
Qy 916 NAGTTAACCTGCTGGCTGTGANTCCCTGTGGATATAAGGGGGMKGGGAAAGGGT 975
Db 841 GAGTTAACCTGCTGGCTGTGANTCCCTGTGGATATAAGGGGGMKGGGAAAGGGT 900
Qy 976 ACTAACCCAGATTTCTGAGCCCTCAGTATGCTCGACATTCATGCTAACCATGACCATGCT 1035
Db 901 ACTAACCCAGATTTCTGAGCCCTCAGTATGCTCGACATTCATGCTAACCATGACCATGCT 960
Qy 1036 TGGATGTCTTACTAGCTGTCTGGGATAGCTTGGAGCACTTACTCAGTGGCTGGTGAAT 1095
Db 961 TGGATGTCTTACTAGCTGTCTGGGATAGCTTGGAGCACTTACTCAGTGGCTGGTGAAT 1020
Qy 1096 GACACCTCAGAGGAATGATGCTTATAGAGAGGAGAGGAG 1137
Db 1021 GACACCTCAGAGGAATGATGCTTATAGAGAGGAGAGAGGAG 1062

RESULT 12
CR594614
LOCUS full-length cDNA clone CS0DJ008YF23 of T cells (Jurkat cell line) 1058 bp mRNA linear HTC 21-JUL-2004
DEFINITION Cot 10-normalized of Homo sapiens (human).
ACCESSION CR594614
VERSION CR594614.1 GI:50475421
KEYWORDS HTC; CNSLT cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1058)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@lifetech.com URL :

```
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1058)
REFERENCE
Genoscope.
AUTHORS
Direct Submission
TITLE
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
Location/Qualifiers
source
1..1058
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ008YF23"
/tissue_type="T cells (Jurkat cell line) Cot
10-normalized"
/plasmid="pCMVSPORT_6"
ORIGIN
Query Match 75.3%; Score 906.8; DB 3; Length 1058;
Best Local Similarity 98.0%; Pred. No. 4e-244; Mismatches 4; Indels 12; Gaps 12;
Matches 1037; Conservative 5;
4Y 98 GCTGCGCTGTGCCCCGCGCCGATTTCTCTCCCGCTGCTT-CCTTCTCTCTGCGCTTC 156
Db 1 GCTGCGCTGTGCCCCGCGCCGATTTCTCTCCCGCTGCTTCTCTCTCTCTGCGCTTC 60
4Y 157 GTGCTCTCTCGCTTCGCGATGGCGGAGCGGCTGCACACAA-GGCGCCCTTCCCTGT 215
Db 61 CTGCTCTCTCGCTTCGCGATGGCGGAGCGGCTGCACCAAGGCGGCTTCCCTGT 120
4Y 216 GATAGGCTCACTTTCTACAGGTCAATCCCAAGCAAGTTCGTC-TGGTGAAGTTCAC 274
Db 121 GATAGGCTCACTTTCTACAGGTCAATCCCAAGCAAGTTCGTC-TGGTGAAGTTCAC 180
4Y 275 ACCAGTACCTTACGCTGAGAGCAGGATGAGTTCAAGCGTCTT-CTGAAACTCGGCT 333
Db 181 ACCAGTACCTTACGCTGAGAGCAGGATGAGTTCAAGCGTCTTCTGTAAGTTCGCT 240
4Y 334 TCCAGCGATGATCTCTTGGTGCGAGGTGGGATCTCAGATTAT-GTGCAAGCTGAAC 392
Db 241 TCCAGCGATGATCTCTTGGTGCGAGGTGGGATCTCAGATTATGTGTGCAAGCTGAAC 300
4Y 393 ATGGAGCTGAGTGAGAAATACAGCTGCACAAAGAGCTACCCA-TCTTCTACCTCTTC 451
Db 301 ATGGAGCTGAGTGAGAAATACAGCTGCACAAAGAGCTACCCAAGTCTTCTACCTCTTC 360
4Y 452 CGGATGGGACTTTGAGAACCCAGTCCCATACCTGCGGCGAGTT-AGTTGGAGCCATC 510
Db 361 CGGATGGGACTTTGAGAACCCAGTCCCATACCTGCGGCGAGTTAAGTTGGAGCCATC 420
4Y 511 CAGCGCTGGCTGAAGGGCAAGGGTCTACTAGGTATGCTCTGG-TGCTGCTGTATAC 569
Db 421 CAGCGCTGGCTGAAGGGCAAGGGTCTACTAGGTATGCTCTGGTTCCTGCTGTATAC 480
4Y 570 GAGCGCTGGCGGGGAGTTCTACAGGGCTCTGTTGTGGAGG-CCGCGAGCCCTCTTG 628
Db 481 GAGCGCTGGCGGGGAGTTCTACAGGGCTCTGTTGTGGAGG-CCGCGAGCCCTCTTG 540
4Y 629 AAGCAGGGGCAAGATTAACCTCTCAAGTGTGAAGGAGCTCAAG-AGTGGGCGGAGCA 687
Db 541 AAGCAGGGGCAAGATTAACCTCTCAAGTGTGAAGGAGCTCAAGAGAGTGGGCGGAGCA 600
4Y 688 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTT-CCAGCATCAGATG 746
Db 601 TACCTGAAGATCATGGGGAAGATCTTAGACCAAGGGGAGCACTTCCAGCATCAGATG 660
4Y 747 ACACGGATCGCCAGGCTGATTGAGAAGAAACAGATGAGTGAC-GGCAGAGGAGGAGCTC 805
```

```
Db 661 ACAGGATCGCAGGCTGATTGAGAGAACAGATGAGTACGGGAAAGAGGAGGCTC 720
4Y 806 CAGAAGAGCTTAAACATCTCAGTCCCTTCAGAGAAAGGGGGCCGAGAAAGAGAGCTG 865
Db 721 CAGAAGAGCTTAAACATCTCAGTCCCTTCAGAGAAAGGGGGCCGAGAAAGAGAGCTG 780
4Y 866 TAAAGAGCTCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGAGTTAACT 925
Db 781 TAAAGAGCTCTGTGTGATTTTCCAGGGTTTGGTGGGGTAGGGAGGGGAGTTAACT 840
4Y 926 GCTGCTGTGANTCCCTTGTGGAATATAGGGGGTMSKGGGAAAGMGGTACTTAACCCAC 985
Db 841 GCTGCTGTGAGTCCCTTGTGGAATATAGGGGGTAGGGGAAAGTGTACTTAACCCAC 900
4Y 996 GATTCTGAGCCCTGAGTATGCTCGACATTCATGATGCTTAACATGACCATCTTGGGATGTCT 1045
Db 901 GATTCTGAGCCCTGAGTATGCTCGACATTCATGATGCTTAACATGACCATCTTGGGATGTCT 960
4Y 1046 CTAGCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGTGTAATGACACCTCAG 1105
Db 961 CTAGCTGTCTGGGATAGCTGGAGCACTTACTCAGGTGGCTGTGTAATGACACCTCAG 1020
4Y 1106 AAGGAATGAGTGTCTATAGAGAGGAGAGGAGTGTACT 1143
Db 1021 AAGGAATGAGTGTCTATAGAGAGGAGAGGAGTGTACT 1058
RESULT 13
BX403041/c
LOCUS
DEFINITION
BX403041 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
Homo sapiens cDNA clone CS0DL010YA17 3-PRIME, mRNA sequence.
ACCESSION
BX403041
VERSION
BX403041.2 GI:46835608
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1098)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
On May 13, 2003 this sequence version replaced gi:30607274.
COMMENT
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6716.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS1AL003ZC09NP1&c=6716.f.
Location/Qualifiers
1..1098
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DL010YA17"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
```

Query Match	75.1%	Score	904.4	DB	5	Length	1098
Best Local Similarity	95.1%	Pred. No.	1.9e-243				
Matches 1041; Conservative	25	Mismatches	16	Indels	13	Gaps	13
QY	83	GAACCCGGGATATGGCTGCGGTGTGCCCGCGCGCATTTCTCTCCCGCTGCTT-CC	141				
DB	1095	GRAMCAGCGDTTATGTGTGCGSTTKTCCCCCSCCCGCGWTTTCTCCCCGCTGCTTCCC	1036				
QY	142	CTTCTCTCGGGCTTCTTGCTCC-TCTCCGCTCGCGCATGGCGGAGCGGCTGCACACCAA	200				
DB	1035	CTTCTCTCGGGCTTCTTGCTCCCTTCTCCSCSTCGCATGGCGGCGGCTGMAACCAA	976				
QY	201	-GGCGCCCTTCCCTCGAATACGGTCACTTTCTTACAAGTCACTTCCAAAAGCAAGTTCGT	259				
DB	975	GGCGCGCCCTTCCCTCGAWACGGTAACTTTCTACAAGTMAATTCCTCCAAAAGCAAGTTCGT	916				
QY	260	C-TGGTGAAGTTGGACACCCAGTACCCCTACGGTGTGAGACGAGTCAAGTTCAGCGTCT	318				
DB	915	CTTGGTGAAGTTCRACACCCAGTACCCCTACGGTGTGAGACGAGTCAAGCGTCT	856				
QY	319	T-CTGMAAACTCGGCTTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATT	377				
DB	855	TGCTGMAAACTCGGCTTTCAGCGATGATCTCTTGGTGGCAGAGGTGGGATCTCAGATT	796				
QY	378	T-GTGCAAGCTGAACATGGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGTACCC	436				
DB	795	TGGTCAAGCTGAACATGAGCTGAGTGAGAAATACAAGCTGGACAAAGAGAGTACCC	736				
QY	437	A-TCTTCTACTCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTCGGGCAGT	495				
DB	735	AGTCTTCTACTCTTTCGGGATGGGACTTTGAGAACCCAGTCCCATACACTCGGGCAGT	676				
QY	496	T-AGGTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACTAGTATGCTCGG	554				
DB	675	TAAGTTTGGAGCCATCCAGCGCTGGCTGAAGGGCAAGGGGTCTACTAGTATGCTCGG	616				
QY	555	-TGCCTGCCTGTATACGACCCCTGGCGGGGAGTTTCATCAGGCGCTCTGGTGTGAGG-	612				
DB	615	TTSCCTGCCGTATACGACCCCTGGCGGGGAGTTTCATCAGGCGCTCTGGTGTGAGGC	556				
QY	613	CCGCCAGGCCCTTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAGAGACTCAGAA	672				
DB	555	CCGCCAGGCCCTTTGAAGCAGGGGCAAGATACTCTCAAGTGTGAGAGACTCAGAA	496				
QY	673	G-AGTGGCGCGACAACTTGAAGATCATGGGGAGATCTTAGACCAAGGGGAGCACTT	731				
DB	495	GAAGTGGCGCGACAACTTGAAGATCATGGGGAGATCTTAGACCAAGGGGAGCACTT	436				
QY	732	-CCAGCATCAGAGATGACCGGATCGCCAGGCTGATTGAGAAGAACAAAGTAGTGAC-G	789				
DB	435	CCCAGCATCAGAGATGACCGGATCGCCAGGCTGATTGAGAGACAAAGTAGTGACGG	376				
QY	790	GCAGAAAGGAGGCTCCAGAAAGCTTTAAACATCCTGACTGCTTCCAGAAAGGGGGC	849				
DB	375	GAAGAAAGGAGGCTCCAGAAAGCTTTAAACATCCTAACTGCTTCCAGAAAGGGGGC	316				
QY	850	CGAGAAAGGAGGCTGTAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGG	909				
DB	315	CGAGAAAGGAGGCTGTAAAGGCTCTCTGTGATTTTCCAGGGTTTGGTGGGGGTAGGG	256				
QY	910	AGGGANAGTTAACTGCTGGCTGTGANTCCCTTGTGGAATATAAGGGGGGMSKGGGAAA	969				
DB	255	AGGGAGAGTTAACTGCTGGCTGTGAGTCCCTTGTGGAATATAAGGGGGGTAGTGGAAA	196				
QY	970	AGMGGTACTAACCCAGATTCTGAGCCCTGAGTATGCTCGGACATTTGATCTCAATGAC	1029				
DB	195	AGTGGTACTAACCCAGATTCTGAGCCCTGAGTATGCTCGGACATTTGATCTCAATGAC	136				
QY	1030	CATGCTTGGGATGTCTCTAGCTGTCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGG	1089				
DB	135	CATGCTTGGGATGTCTCTAGCTGTCTGGGGATAGCTGGAGCACTTACTCAGGTGGCTGG	76				
QY	1090	TGAAATGACACCTCAGAAAGGAATGAGTGTCTATAGAGAGGAGAGAGGAGTGTACTGCCAC	1149				

```

Db      75  TGAATGACACCTCAGAGGAATGAGTGTATAGAGGAGGAGAGGTACTGTCAG 16
      1150 GTCTTTGACAGATGT 1164
      15  GTCTTIVACAGGTG 1

RESULT 14
AL561169      1093 bp      mRNA      linear      EST 02-APR-2004
LOCUS      AL561169 Homo sapiens B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED
DEFINITION      Homo sapiens cDNA clone CS0DL001YK02 5-PRIME, mRNA sequence.
ACCESSION      AL561169
VERSION      AL561169.3 GI:46186530
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1093)
AUTHORS      Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL      Unpublished (2001)
COMMENT      On Feb 15, 2001 this sequence version replaced gi:31285298.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
6716.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DL001BF01Q1P1&c=6716.f.
Location/Qualifiers
1. 1093
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cloned="CS0DL001YK02"
/cell_type="B CELLS (RAMOS CELL LINE) COT 25-NORMALIZED"
/clone_line="RAMOS CELL LINE"
/clone_lib="Homo sapiens B CELLS (RAMOS CELL LINE) COT
25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match      70.6%; Score 850.4; DB 1; Length 1093;
Best Local Similarity 97.3%; Pred.No. 3.3e-228;
Matches 991; Conservative 8; Mismatches 6; Indels 13; Gaps 13;

Qy      88  CGCGCATATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCGTGT-CCCTCT 146
      1  CGCGCATATGGCTGCGCTGTGCCCCGCGCGCATTTCTCTCCCGCGTGTCTCCCTCT 60
      147  CTTGGGCTTCTGCTCTCTCCGCTCCGATGGCGGAGCGGCTGCACACCAA-GCGGC 205
      61  CTTGGGCTTCTGCTCTCTCCGCTCCGATGGCGGAGCGGCTGCACACCAAAGGGCGC 120
      206  CTTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTCGTC-TCGT 264
      121  CTTTCCCTCGATACGGTCACTTTCTACAAGGTCAATCCCAAAAGCAAGTTCGTTGGT 180
      265  GAAGTTTCGACACCCAGTACCCCTACGCTGAGAGACAGATGAGTTCAAGCGTCTTT-CTGA 323
      181  GAAGTTTCGACACCCAGTACCCCTACGCTGAGAGACAGATGAGTTCAAGCGTCTTGTGA 240

```



```
Qy 781 TGAAGTAC-GCCAGAAGGAGGAGCTCCAGAAGAGCTTAAACATCCTGACTGCCTTCCAGA 839
Db ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 409 TGAAGTACGGGAAGAGGAGGAGCTCCAGAAGAGCTTAAACATCCTGACTGCCTTCCAGA 350
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 840 AGAAGGGGGCCGAGAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGT 899
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 349 AGAAGGGGGCCGAGAAGAGGAGCTGTAAAAGGCTGTCTGTGATTTTCCAGGGTTTGGT 290
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 900 GGGGTAGGGAGGGGANAGTTAACTGCTGCTGTGANTCCCTTGTGGAATATAAGGGGG 959
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 289 GGGGTAGGGAGGGGAGHNTAACCTGCTGCTGTGAGTCCCTTGTGGAATATAAGGGGG 230
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 960 YMSKGGGAAAAGWGTACTAACCCACGATTTCTGAGCCCTGAGTATGCCCTGGACATTGATG 1019
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 229 TAGTGGGAAAAGTGGTACTAACCCACGATTTCTGAGCCCTGAGTATGCCCTGGACATTGATG 170
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1020 CTAACATGACCATGCTTGGGATGTCTCTAGCTGGTCTGGGGATAGCTGGAGCACTTACTC 1079
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 169 CTAACATGACCATGCTTGGGATGTCTCTAGCTGGTCTGGGGATAGCTGGAGBACTTACTC 110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1080 AGGTGGCTGTGAATGACACCTCAGAGGAATGAGTGTCTATAGAGAGGAGAGAGGAGTG 1139
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 109 AGGTGGCTGTGAATGACACCTCAGAGGAATGAGTGTCTATAGAGAGGAGAGAGGAGTG 50
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1140 TACTGCCCAGGT--CTTTGACAGATGTAAATTTCTCAATTAAGTTT 1186
Db ||||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 49 TACTGCTNCAGKTCCTTTTATCAGGTGTAAATTCNNTCDWGTTAAHTTT 1
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: July 3, 2005, 16:39:54
Job time : 3817.04 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 07:04:57 ; Search time 10073.7 Seconds
(without alignments)
12038.564 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTGACAGG.....GGACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gse1:*

9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1028.6	32.3	2331	3 AK030249	Mus musculus
2	762.2	23.9	841	6 CD644465	AGENCOURT
3	647	20.3	651	5 BX119280	Medline
4	615.6	19.3	618	5 BX955300	DKFZp781N
5	590	18.5	595	4 BF944766	PM3-NN117
6	581.4	18.2	583	5 BP336988	BP336988
7	471.8	14.8	569	2 AW753609	RC1-CT026
8	470.8	14.8	895	9 CR094054	Forward
9	466	14.6	479	1 A1380386	Forward
10	453.2	14.2	805	6 CB522302	UI-N-GH0-
11	424.4	13.3	692	2 BB617224	BB617224
12	418.8	13.1	768	2 BP237807	BP237807
13	395.2	12.4	473	7 W05407	zab3c12.1
14	394	12.4	411	4 BG007233	IL5-GN023
15	384	12.1	640	4 BM834501	K-EST0109
16	377.4	11.8	594	7 CV409050	RC1-HT026
17	376.2	11.8	636	6 CB457791	715904.MA
18	369.2	11.6	519	1 AI025334	ov74905.x
19	369	11.6	399	2 BB819616	RC2-BN033
20	368.6	11.6	400	5 BQ367562	IL5-GN023
21	365.4	11.5	368	4 BM146822	TCAAP1E74
22	350.8	11.0	412	1 AA069107	zm10b10.x
23	347.4	10.9	703	4 B1185557	UNL-P-PN-
24	346.2	10.9	353	4 BG013874	IL5-GN023

C 25	344.6	10.8	662	6	CB460210	CB460210
C 26	343.2	10.8	440	7	N70546	zab3c12.81
C 27	318	10.0	674	6	BY748714	BY748714
C 28	309.4	9.7	667	8	AZ939266	AZ939266
C 29	301.8	9.5	305	4	BG982616	IL5-CN006
C 30	300	9.4	352	4	BF963063	PM3-NN117
C 31	299	9.4	328	2	BE818091	CM2-BN027
C 32	294.4	9.2	571	7	CR753532	DKF2P469L
C 33	284.2	8.9	327	2	BF782572	602106595
C 34	282	8.9	327	2	BE817937	CM2-BN027
C 35	278.2	8.7	848	1	AJ743843	AJ743843
C 36	264.4	8.3	313	3	BF803316	RC1-CI011
C 37	262.6	8.2	284	2	BE819568	RC2-BN033
C 38	260.4	8.2	308	4	BG012719	IL5-GN023
C 39	258	8.1	308	2	BE818055	CM2-BN027
C 40	256.4	8.0	341	2	BF329627	CM2-BN027
C 41	254	8.0	332	4	BG978004	PM1-CI015
C 42	251.4	7.9	254	1	AU185193	AU185193
C 43	248.6	7.8	339	4	BG012812	IL5-GN023
C 44	247.2	7.8	749	1	AJ743855	AJ743855
C 45	246.6	7.7	262	2	AW380095	QV3-HT026

ALIGNMENTS

RESULT 1	AK030249	AK030249	2331 bp	musRNA	linear	HTC 03-APR-2004
LOCUS	AK030249	Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:4933431C08 product:weakly similar to ZINC FINGER PROTEIN 217 [Homo sapiens], full insert sequence.				
DEFINITION	AK030249	HTC; CAP trapper.				
ACCESSION	AK030249	Mus musculus (house mouse)				
VERSION	AK030249.1	GI:26081731				
KEYWORDS	HTC; CAP trapper.					
SOURCE	Mus musculus					
ORGANISM	Mus musculus					
REFERENCE	1	Carninci, P. and Hayashizaki, Y.				
AUTHORS	1	High-efficiency full-length cDNA cloning				
TITLE	1	Mech. Enzymol. 303, 19-44 (1999)				
JOURNAL	1	99279253				
MEDLINE	1	10349636				
PUBMED	1	10349636				
REFERENCE	2	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
AUTHORS	2	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
TITLE	2	Genome Res. 10 (10), 1617-1630 (2000)				
JOURNAL	2	20499374				
MEDLINE	2	11042159				
PUBMED	2	11042159				
REFERENCE	3	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.				
AUTHORS	3	RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer				
TITLE	3	Genome Res. 10 (11), 1757-1771 (2000)				
JOURNAL	3	20530913				
MEDLINE	3	11076861				
PUBMED	3	11076861				
REFERENCE	4	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.				
AUTHORS	4	Functional annotation of a full-length mouse cDNA collection				
TITLE	4	Nature 409, 685-690 (2001)				
JOURNAL	4					
MEDLINE	4					
PUBMED	4					
REFERENCE	5	The FANTOM Consortium and the RIKEN Genome Exploration Research				

Db 241 CACCAGGCAACAGCAATCTGAGATGTTTCTTAAACACAGTGTCTCCCTGACCGGATAA 300
Qy 2559 GACAAAAGAGCCGAGACAAAATGAAACCTTCCAGTAGCTCTTCTCAGGCCACCT 2618
Db 301 GACAAAAGAGCCGAGACAAAATGAAACCTTCCAGTAGCTCTTCTCAGGCCACCT 360
Qy 2619 CGGACAGAGTAACATCAATGCTTCATCGACTACCCGCGCAAGAACGACAGCCCGTGGGC 2678
Db 361 CGGACAGAGTAACATCAATGCTTCATCGACTACCCGCGCAAGAACGACAGCCCGTGGGC 420
Qy 2679 ACCTCCGGAAGAGACTATTTCTGTAATCGAGTGGCCAGCAATATCTGAGCAGAAATTTGG 2738
Db 421 ACCTCCGGAAGAGACTATTTCTGTAATCGAGTGGCCAGCAATATCTGAGCAGAAATTTGG 480
Qy 2739 TGAGCCCTTCCAAAAGAGACTGAAAGTCCAGCGTGGTTCCTTCAGCTTGACACGCGCG 2798
Db 481 TGAGCCCTTCCAAAAGAGACTGAAAGTCCAGCGTGGTTCCTTCAGCTTGACACGCGCG 540
Qy 2799 GGCCAAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATC 2858
Db 541 GGCCAAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATC 600
Qy 2859 ACTGTTACCGCAGAGCTGTGTATTCCTGTCGAGCGCTGCTCCCAAACCAAGGTTCT 2918
Db 601 ACTGTTACCGCAGAGCTGTGTATTCCTGTCGAGCGCTGCTCCCAAACCAAGGTTCT 660
Qy 2919 GAGCTCCAGCAGCTCGATTTCCAAATGCTGCTGCTTCCAGAGCCCTATGTCGCTC 2978
Db 661 GAGCTCCAGCAGCTCGATTTCC -AATGTCTGACTGTTTCCAGAAAGCCTATGTCGCTC 719
Qy 2979 CGGGCCACTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGAGCTTAGAAG 3038
Db 720 CGGGCCACTTTACACTTGTGTGCTGCTGTGTAGTCCAGCATCCAGCTCGAGCTTAGAAG 779

RESULT 3
BX119280
LOCUS
DEFINITION BX119280 Soares_fetal_lung_NbHL19W Homo sapiens cDNA clone
ACCESSION IMAGp998F23669 ; IMAGE:299158, mRNA sequence.
VERSION BX119280
KEYWORDS BX119280.1 GI:27882834
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ebert L., Heil O., Hennig S., Neubert P., Partsch E., Peters M.,
Radelof U., Schneider D. and Korn B.
Human Unigeneset - RZPD3
Unpublished (2003)
Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998F23669,
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
http://www.rzpd.de/CloneCards/cgi-
bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACACGAGAACAGCTATGAC.
Location/Qualifiers
1..651
/organism="Homo sapiens"
/mol_type="mRNA"

FEATURES
source

/db xref="taxon:9606"
/clone="IMAGp998F23669 ; IMAGE:299158"
/dev stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung_NbHL19W"
/notes="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTCTTTTCTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M.Patima Bonaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NbHL19W."
ORIGIN
Query Match 20.3%; Score 647; DB 5; Length 651;
Best Local Similarity 99.5%; Pred. No. 6.3e-176;
Matches 647; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1107 TAGCAAGAGAGCCCACTCCTCGAGTGGCGCAAGCTTTTCAGAACCTTACCACCA 1166
Db 1 TAGCAAGAGAGCCCACTCCTCGAGTGGCGCAAGCTTTTCAGAACCTTACCACCA 60
Qy 1167 GCTGCTTCTGACCGGAGGCGCGGAGCCGCGGAGCGCGGAGTCCGCCAC 1226
Db 61 GCTGCTTCTGACCGGAGGCGCGGAGCCGCGGAGCGCGGAGTCCGCCAC 120
Qy 1227 CATGCTGTGACCGGAGGCGCGGAGCGTGTCTCTGACCTCGCGGCCCTCTGGA 1286
Db 121 CATGCTGTGACCGGAGGCGCGGAGCGTGTCTCTGACCTCGCGGCCCTCTGGA 180
Qy 1287 TGAATAATGGAGCGCTGATCGAGGGGAAGTGTGTTCTGAAGACGGATCTGAGGATGGCT 1346
Db 181 TGAATAATGGAGCGCTGATCGAGGGGAAGTGTGTTCTGAAGACGGATCTGAGGATGGCT 240
Qy 1347 TCCCGAAGGAATCCATCTGGATAAAATGATGAGGAGGAGAAATAAACAATCTTACATC 1406
Db 241 TCCCGAAGGAATCCATCTGGATAAAATGATGAGGAGGAGAAATAAACAATCTTACATC 300
Qy 1407 TTCAAGAGAGTGTAGTTATTGTGAAAGTCTTCCGTTCAAATTTATTCCTCAATATTTCA 1466
Db 301 TTCAAGAGAGTGTAGTTATTGTGAAAGTCTTCCGTTCAAATTTATTCCTCAATATTTCA 360
Qy 1467 TCTCAGAACGCATACAGGTGAAAACCATACAAATGAAATTTTGTGAATATGCTGCAGC 1526
Db 361 TCTCAGAACGCATACAGGTGAAAACCATACAAATGAAATTTTGTGAATATGCTGCAGC 420
Qy 1527 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAGAAACAAACCGATGT 1586
Db 421 CCAGAAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAGGAGAAACAAACCGATGT 480
Qy 1587 TGCTGCTGAAGTCAAGAACCGATGTTAAATAATCAGGACATCAAGATGCACTATTAAACCGC 1646
Db 481 TGCTGCTGAAGTCAAGAACCGATGTTAAATAATCAGGACATCAAGATGCACTATTAAACCGN 540
Qy 1647 TGACAGTGGCGAAACCAAAAATTTGAAAAGATTTTTTGATGCTGCCAAGATGTTACAGG 1706
Db 541 TGACAGTGGCGAAACCAAAAATTTGAAAAGATTTTTTGATGCTGCCAAGATGTTACAGG 600
Qy 1707 CAGTCCACCTGCAAGCAGCTTAAGGAGATCCCTTCTGTTTTTTCAGAAATG 1756
Db 601 CAGTCCACCTGCAAGCAGCTTAAGGAGATCCCTTCTGTTTTTTCANAAATG 650
RESULT 4
BX955300
LOCUS
DEFINITION DXFzp781N0270_r1 781 (synonym: hlccc4) Homo sapiens cDNA clone
DXFzp781N0270 5', mRNA sequence.

```

ACCESSION   BX955300
VERSION     BX955300.1  GI:43437635
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 618)
AUTHORS    Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
            Pobo, G., Han, M. and Wiemann, S.
TITLE      EST (Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., et al.)
JOURNAL     Unpublished (2003)
COMMENT     Contact: MIPS
            MIPS
            Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
            researched by AGOWA (Berlin/Germany) within the cDNA sequencing
            consortium of the German Genome Project.
            No sl sequence available.
            This clone (DKFZp781N0270) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES    source
            1..618
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="DKFZp781N0270"
            /dev_stages="adult"
            /lab_host="DH10B"
            /clone_lib="781 (synonym: hlcc4)"
            /note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
            cDNA-collection"
ORIGIN
Query Match      19.3%; Score 615.6; DB 5; Length 618;
Best Local Similarity 99.0%; Pred. No. 8.3e-167;
Matches 612; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
Qy 58 GGGCCAGAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCTTGTGCA 117
Db 1 GGGCCAGAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCTTGTGCA 60
Qy 118 ATGAAAGGACCGCTGTGTTCCATTCGAGCTACACAAAGAAAAATGTTCATCCAAATC 177
Db 61 ATGAAAGGACCGCTGTGTTCCATTCGAGCTACACAAAGAAAAATGTTCATCCAAATC 120
Qy 178 GAGGGGTATGCCCTTGGATTGCGATGTTCTGCAGCCAGACCTTCACACATTCAGAGAC 237
Db 121 GAGGGGTATGCCCTTGGATTGCGATGTTCTGCAGCCAGACCTTCACACATTCAGAGAC 180
Qy 238 CTTAATAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACACGAGTCTTCTCGG 297
Db 181 CTTAATAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACACGAGTCTTCTCGG 240
Qy 298 GTTGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTGCAGACAGACCTCCCAAG 357
Db 241 GTTGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTGCAGACAGACCTCCCAAG 300
Qy 358 GAAAGATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCACACATTTAGAGTC 417
Db 301 GAAAGATTCAGAGAAATGAATTTAGCTGTGAGGTATGTGGGCACACATTTAGAGTC 360
Qy 418 GCTTTTGTATGTTGAGATCCATGAGACACACAAAGATCTTTCACCTTCGGGTGAAC 477
Db 361 GCTTTTGTATGTTGAGATCCATGAGACACACAAAGATCTTTCACCTTCGGGTGAAC 420
Qy 478 ATGTGCGGAAGAGMTTSSRSAGCCTTGGTTTCTTAAATAATCATCATCGGACACATAAT 537
Db 421 ATGTGCGGAAGAGMTTCAAGGAGCCTTGGTTTCTTAAATAATCATCATCGGACACATAAT 480

```

```

Qy 538 GGCAAAATCGGGCCAGAAAGCAAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATC 597
Db 481 GGCAAAATCGGGCCAGAAAGCAAACTGCAGCAAGGCTTTGGAGAGTAGTCCAGCAACGATC 540
Qy 598 AACGAGGTGCTGCAGGTGCACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCATG 657
Db 541 AACGAGGTGCTGCAGGTGCACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCATG 600
Qy 658 GTTTGTGGCTTCCTATT 675
Db 601 GTTTGTGGCTTCCTATT 618

RESULT 5
BF944766
LOCUS      BF944766
DEFINITION PM3-NN1174-181000-001-c08 NN1174 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF944766
VERSION    BF944766.1  GI:12362041
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
REFERENCE   1 (bases 1 to 595)
AUTHORS     Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
            O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE     2026663
PUBMED      10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM3&t2=PM3-NN1174-
            181000-001-c08&t3=2000-10-18&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 585.
FEATURES    source
            1..595
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /dev_stages="adult"
            /clone_lib="NN1174"
            /note="Organ: nervous normal; Vector: puc18; Site_1: SmaI;
            Site_2: SmaI; A mini-library was made by cloning products
            derived from ORESTES PCR (U.S. Letters Patent application
            No. 196,716 - Ludwig Institute for Cancer Research)
            profiles into the pUC 18 vector. Reverse transcription of
            tissue mRNA and cDNA amplification were performed under
            low stringency conditions."
ORIGIN
Query Match      18.5%; Score 590; DB 4; Length 595;
Best Local Similarity 98.8%; Pred. No. 2.3e-159;
Matches 587; Conservative 6; Mismatches 1; Indels 0; Gaps 0;
Qy 297 GGTGGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTGCAGACAGACCTCCCAA 356
Db 1 GGTGGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTGCAGACAGACCTCCCAA 60

```

```
Qy 357 GGAAGAAGATTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGCGAGACATTTAGAGT 416
Db 61 GGAAGAAGATTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGCGAGACATTTAGAGT 120
Qy 417 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCCTTTCATTACGGGTGTAA 476
Db 121 CGCTTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCCTTTCATTACGGGTGTAA 180
Qy 477 CATGTGCGGAAGAAGMTTSSRSSAGCCTTGGTTTCTTAAATAATCACATGCGGACACATAA 536
Db 181 CATGTGCGGAAGAAGATTCAAGGAGCCTTGGTTTCTTAAATAATCACATGCGGACACATAA 240
Qy 537 TGGCAAAATCGGGGGCGAGAACAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAAAACGAT 596
Db 241 TGGCAAAATCGGGGGCGAGAACAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAAAACGAT 300
Qy 597 CAACGAGTGCTCGAGTGCACGCGGCGGAGAGCATTCCTCTCTTACAAAATCTGCAT 656
Db 301 CAACGAGTGCTCGAGTGCACGCGGCGGAGAGCATTCCTCTCTTACAAAATCTGCAT 360
Qy 657 GGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 716
Db 361 GGTTTGTGGCTTCCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACAC 420
Qy 717 CAAAAAAACTGCTTTTCGGTACGAGCGCGGAGAGACTCTCCACAAAGGAGGAATGCC 776
Db 421 CAAAAAAACTGCTTTTCGGTACGAGCGCGGAGAGACTCTCCACAAAGGAGGAATGCC 480
Qy 777 GTCTCTGAGGAGGAGACTTCCTGCGAGTTGTTTCAACTTGAGACAAAATCTACCCGTGAAC 836
Db 481 GTCTCTGAGGAGGAGACTTCCTGCGAGTTGTTTCAACTTGAGACAAAATCTACCCGTGAAC 540
Qy 837 GGGGAAGAAGCTGTGCAGATCATCCCTCAGCTCGATCCGTTCCACCACCTTCCA 890
Db 541 GGGGAAGAAGCTGTGCAGATCATCCCTCAGCTCGATCCGTTCCACCACCTTCAA 594

RESULT 6
LOCUS BP336988 Sugano cDNA library, coronary artery smooth muscle cell EST 17-SEP-2004
DEFINITION Homo sapiens cDNA clone SMR07185, mRNA sequence.
ACCESSION BP336988
VERSION BP336988.1 GI:52266581
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 583)
Suzuki Y., Yamaehita, R., Shiota, M., Sakakibara, Y., Chiba, J.,
Mizushima-Sugano, J., Nakai, K. and Sugano, S.
Sequence comparison of human and mouse genes reveals a homologous
block structure in the promoter regions
Genome Res. 14 (9), 1711-1718 (2004)
Contact: Yutaka Suzuki
Department of Virology
Institute of Medical Science, University of Tokyo
4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
Email: yusuzuki@ims.u-tokyo.ac.jp.
Location/Qualifiers
FEATURES
1..583
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="SMR07185"
/tissue_type="coronary artery"
/cell_type="smooth muscle cell"
/clone_lib="Sugano cDNA library, coronary artery smooth
muscle cell"

ORIGIN
```

```
Query Match 18.2%; Score 581.4; DB 5; Length 583;
Best Local Similarity 99.8%; Pred. No. 7,1e-157;
Matches 582; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1626 TGRAGATGCACTATTACCGCTGACAGTGGCGCAACCAAAATTTGAAAAGATTTTGA 1685
Db 1 TGAAGATGCACTATTACCGCTGACAGTGGCGCAACCAAAATTTGAAAAGATTTTGA 60
Qy 1686 TGGTGCCAAAGATGTTACAGGAGCTCCACCTGCAAAAGCAGCTTAAGGAGATGSCCTTCTGT 1745
Db 61 TGGTGCCAAAGATGTTACAGGAGCTCCACCTGCAAAAGCAGCTTAAGGAGATGSCCTTCTGT 120
Qy 1746 TTTTCAGAAATGTTGGGCGAGCGCTGCTCTCAACAGCACACAAAGATATCTAGGATTT 1805
Db 121 TTTTCAGAAATGTTGGGCGAGCGCTGCTCTCAACAGCACACAAAGATATCTAGGATTT 180
Qy 1806 CCATAAAATGCAGCTGATGACAGTGTGATAAAGTGAATAAAACCTCACCCCTGCTTA 1865
Db 181 CCATAAAATGCAGCTGATGACAGTGTGATAAAGTGAATAAAACCTCACCCCTGCTTA 240
Qy 1866 CCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATCTGTAG 1925
Db 241 CCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATACCTCATCTGTAG 300
Qy 1926 AACCAAGCGGATGTTACTCTCTCCGATGGCAGTACCAACCAACCTTGAAGTTAG 1985
Db 301 AACCAAGCGGATGTTACTCTCTCCGATGGCAGTACCAACCAACCTTGAAGTTAG 360
Qy 1986 CCCCAGAGAGAACAAACGGAGACCGCAGTGTGAGATACAGGCAAGCTGAGTTG 2045
Db 361 CCCCAGAGAGAACAAACGGAGACCGCAGTGTGAGATACAGGCAAGCTGAGTTG 420
Qy 2046 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAAATGCCCGCAATTTCTTT 2105
Db 421 TCACGAAAAACCTTTAAATTTATCCGTGGGGGCTCTTCACAAATGCCCGCAATTTCTTT 480
Qy 2106 GAGTAAAGTTTGATCCAAAGTATCACGTGTCATTTTGTACCTTCAAGACATTTTATCC 2165
Db 481 GAGTAAAGTTTGATTCCAAGTATCACGTGTCATTTTGTACCTTCAAGACATTTTATCC 540
Qy 2166 AGAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAT 2208
Db 541 AGAAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAT 583

RESULT 7
LOCUS AW753609/c 569 bp mRNA linear EST 28-APR-2000
DEFINITION RCI-CT0268-060100-013-h08 CT0268 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW753609
VERSION AW753609.1 GI:7668541
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 569)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
```

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC1&t2=RC1-CT0268-060100-013-h08&t3=2000-01-06&t4=1>)
 Seq primer: puc 18 forward
 High quality sequence start: 49
 High quality sequence stop: 67.
 Location/Qualifiers
 1. 569
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="CT0268"
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the PUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN	Query Match	14.8%	Score 471.8	DB 2	Length 569
	Best Local Similarity	92.1%	Pred. No. 5.2e-125		
	Matches 524	Conservative 5	Mismatches 37	Indels 3	Gaps 3
QY	113	TGTCATGAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCATCC	172		
DB	569	TGTCATGAAGGACCGCTGTTGTTCCATTCGAGCGACACAAGAATAAGATGTCATCC	510		
QY	173	AAATCGAGGGGTATATGCCCTTGATTTGCATGTTCTGCA-GCCAGACCTTCACACATTCA	231		
DB	509	CATCGAGGGGTATATGCCCTTGATGCGCATGTTCTGCATGCCAGACTCGACACATTCA	450		
QY	232	GAAGACCTTAATAACAATGTCCTTAATGGAAACACCGGCGCTACCC--TCGTGAAACGACGAT	290		
DB	449	GAAGACCTGAATCCACATGTCCTGATGCGACACCGGCGTACCCCTTCGTGAAACCGACGAT	390		
QY	291	TC TTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTCGGAACAGAAACC	350		
DB	389	TC TTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGATATAAAGTCAAGTCGGAACAGAAACC	330		
QY	351	TCCCAAGGAAAAAGAAATTCGAAGGAAAAATGAATTTAGCTGTGAGTATGTGGCGCAGACATT	410		
DB	329	TCCCAAGGAAAAAGAAATGCGAAGGCAAAATGAATTTAGCTGTGAGTATGTGGCGCAGACATT	270		
QY	411	TAGAGTCGCTTTTGATGCTTGAGATCCACATGAGNACACACAAGAGATCTTTCACCTACGG	470		
DB	269	TAGAGTCGCTTTTGATGCTTGAGATCCACATGAGNACACACAAGAGATTCGGTCTACTGACGG	210		
QY	471	GTGTAA CATGTGCGGAAGAGMTTSSRSAGCCTTGCTTTCTTAAAAATCACATGCGGAC	530		
DB	209	GTGTGACATGTGCGGTAGAAGATTTAAGGAGCCTTGCTTTCTTAAATACATGCGGAC	150		
QY	531	ACATAATGGCAAAATCGGGGGCGCAGAAGCAAATCTGCAGCAAGGGCTTGAGAGTAGTTCAGC	590		
DB	149	ACATAATGGCAAAATCGGGGGCGCAGAATCAATCTGCAGCAAGGGCTTGAGAGTAGTTCAGC	90		
QY	591	AACGATCAAAGAGTTCGTCCAGGTGACCGGGCGCAGAGCATCTCCTCTCC--TTACAAAA	649		
DB	89	CACGATCAAAGAGTTCGTCCAGGTGACCGGGCGCAGAGCATCTCCTCTCC--TTACAAAA	30		
QY	650	TC TCGATGGTTTGGGCTTCCTATTTCGA	678		
DB	29	TC TCGATGGTTTGGGCTTCCTATTTCGA	1		

RESULT 8
CR094054

LOCUS	CR094054	895 bp	DNA	linear	GSS 05-JUL-2004
DEFINITION	Forward strand read from insert in 5'HPRT insertion targeting and chromosome engineering clone MHPN19i16, genomic survey sequence.				
ACCESSION	CR094054				
VERSION	CR094054.1	GI:49827883			
KEYWORDS	GSS; genome survey sequence; MICER.				
SOURCE	Mus musculus	(house mouse)			
ORGANISM	Mus musculus				
REFERENCE	Mammalia; Eutheria; Rodentia; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Muridae; Murinae; Mus.				
AUTHORS	1 (bases 1 to 895) Adams,D.J., Biggs,P.J., Cox,A.V., Davies,R.M., van der Weyden,L., Jonkers,J., Smith,J., Plumb,R.W., Taylor,R.G., Nishijima,I., Yu,Y., Rogers,J. and Bradley,A.				
TITLE	Direct Submission				
JOURNAL	Submitted (20-FEB-2004) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. http://www.sanger.ac.uk/MICER				
FEATURES	location/Qualifiers				
source	1..895				
ORIGIN	/organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" /clone="MHPN19i16" /clone_lib="MHPN"				
Query Match	14.8%;	Score 470.8;	DB 9;	Length 895;	
Best Local Similarity	72.0%;	Pred. No. 1.2e-124;			
Matches 647;	Conservative 6;	Mismatches 239;	Indels 7;	Gaps 3;	
Qy	13	GTGACGAGAAACATGCCAACTCAATCCCTCTTAATGTATCATGGATGGCCGAGAGTGATT	72		
Db	1	GTGTTGGCAGCATCGACCCAGTCCCTCTCTGTGT-CATGGACGGCCGGAAGTCCCTC	59		
Qy	73	GGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCTTGTCAATGAAGGGACCGCT	132		
Db	60	AGCAGCTCTCTAGGCTCCCAAGTGGAGTGGATGTCTGTGCCATAAAGGGCCGGT	119		
Qy	133	GTGTGTTCCATTCCGAGCTACACAAGAAAAAATGTCTATCCAAATCGAGGGGTATATGCC	192		
Db	120	GCAGTCCCTTCCGAGTGTCTCAGGAGAGAGAGCATGGCCGTGGCAGGGGCCACATGCC	179		
Qy	193	TTGGATTGCATGTTCTGCAGCCACAGCCTTCCACACATTCAGAAGACCTTAAATAACATGTC	252		
Db	180	CTGGATTGCATGTTCTGCAGCCAGGTCTTCTCTCAGGCGGAGGATCTCACTCAGCAGTG	239		
Qy	253	TTAATGGAACACCGGCTTACCTCTGTGAAACAGCAGTTCCTCGGTTGTAAGCAGAGTAT	312		
Db	240	CTCTCGCAGCACCGGCCACCTCTCGAGCCAGCTGTCTCTGCTGTGTGGAGGCGCAGTAC	299		
Qy	313	CTCAGTCCGTTGATTAAGTCAAGTGGCAGACAGAACCTCCCAAGGNAAGAAATGCCAAG	372		
Db	300	CTAAGTCCCTTGATAAAGCTCTGGAGCCCAACAGACAGCATTTGGAGAAAGTGCGCA	359		
Qy	373	GA---AAATGAA'TTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTATGTT	429		
Db	360	GACCCGAGGAGTTGAGCTGTGATGTGTGTGGCGAGACATTCACAGTGGCTTTTGTATGTT	419		
Qy	430	GAGATCCACATGAGAAACACAAAGATTTCTTTCAC'TTACGGGTGTAACTATGTCGGGAAGA	489		
Db	420	GAGAGCCACATGAAGAAGCATAGGACTCCTTCACTATGGGTGCAGCATGTGCGGGAGG	479		
Qy	490	AGMTTSRRSAGCCTTGGTTTCTTAAATATCATATCGGNCACATATATGGCAATCGGG	549		
Db	480	AGATTCAAGGAGCGGTGGTTCCTTGAAGAAACCATCGGACACACAATATGGCAATCTGGC	539		
Qy	550	GCCAGAAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTC	609		
Db	540	ACGAGGNCAGCTTCAGCAAGGCATGGAG---AGTCCAGTCAACCATCAATGAAGTGCTC	596		
Qy	610	CAGGTGCACGCGCCGAGAGCATCTCTCTCTTACAAATCTGCATGTTGTGCTTC	669		
Db	597	CACCGCACGCCCTGGGAGCATCTCACGCCCTACAAGATCTGCATGTTCTCGGCTTC	656		

1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATCTGAT. This library was created for the University Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 14.2%; Score 453.2; DB 6; Length 805;
Best Local Similarity 78.9%; Pred. No. 1.5e-119;
Matches 633; Conservative 0; Mismatches 145; Indels 24; Gaps 7;

QY 1030 CTCTCGCAAGAGAGAGAGTGTCAAACTCCACGGCGAGCGCCCTCCGTGACGCG 1089
DB 1 CTCTCCCAAGCAAGAGAGAGCTAGACATGCTAATAGTGAAGTGCCTTCTGGGATAGT 60

QY 1090 GATCCCAAGTTACCCAGTAGCAGGAGAGCCCACTCACTGCTCCGAGTGGCGCAAGCT 1149
DB 61 GACCCCAAGTTCTCCAGTAGCAGGAGAGCCCACTGTTCTGAGTGACGCAAGCC 120

QY 1150 TTCAGAACCTACACAGCTGTCTTGCACTCCAGGGTCCACAGAGACCGGAGGCC 1209
DB 121 TTCAGGACATACACAGCTGTCTTGCACTCCAGGGTCCACAGAGACCGGAGGACT 180

QY 1210 GCGCGGAGTCCGCCACCATGCTCTGCGAGGAGCGCGGACGCTTCTCTGAC 1269
DB 181 GATGCCCTGTCCACCATGCTGTGAGTGAAGCGAGCTGCGACCTGCTCCCGACAC 240

QY 1270 CTCGCGCGCCCTCTCGATGAAATGGAGCGGTGAGTCGAGGGGAAGTGGTCTCAAGAC 1329
DB 241 CTCAGCACCACTCTGGAGACAGTGGCGCGGACCGA---GAGGGGGCTCTGAAGAC 297

QY 1330 GATCTGAGGATGGGCTTCCGAGAGATTCATCTGATATAAAATGATGAGGAGAAA 1389
DB 298 GGGTCTGAGGATGGACTCCCTGACGGGTCCATTTGGATATAAAATGATGAGGAGAAA 357

QY 1390 ATAAACATCTTACATCTCAGAGAGTGTAGTTATTGGGAAGCTTTTCGGTTCAAT 1449
DB 358 GCGAAGCCCCCTCCGCTCTCGAGAGAGTGTAGTTACTGTGGCAAGTCTTTCGGTTCAAC 417

QY 1450 TATTACCTCAATATTCATCTCAGACGATACAGTGTGAAACCAACATACAAATGTCAATTT 1509
DB 418 TATTACCTCAATATTCATCTCAGACGATACAGTGTGAAACCAACATACAAATGTCAATTC 477

QY 1510 TGTGAATATGTGCGAGCCGAGAGACATCTCTGAGGTATCAC-TTGGAGAGACATCAAA 1568
DB 478 TGTGAGTATGCGCGAGCCGAGAGACATCTCTGAGGTATCAC-TTGGAGAGACATCAAA 537

QY 1569 GGAAGAAACAAAC---CGATGTTGCTGCTGAAGTCAAGACGATGTGTAATAATCAGGACAC 1625
DB 538 AGACAAGCAGCGCGGTGCTGCGCGCTGAGTCCAAAGAGTGAAGCCGCGAGCGAGCC 597

QY 1626 TGAAGATGCATATTAA---CGCTGACAGTGCAGCAACCAAAATTTGAAAAGATTTT 1682
DB 598 GCAGGATGCGCTTAAACGCTGCTGACAGTGGCAGACCAAAATTTAAAGAGATTTCT 657

QY 1683 TGATGGTGCCAAAGATGTTTACAGGCGAGTCCACCTGCAAGACGCTTAAAGGAGATCCCTTC 1742
DB 658 TGATGGTGCCAAAGATGTTA-AGGAGAGCCACCTGCGAAGCAGCTTAAAGGAGATG-CTTC 715

QY 1743 TGTTTTTCAAGATGTTCTGGGCGAGCGCTGCTCTCACAGCAGCAACAAAGATCTCAGGA 1802
DB 716 TGTTCTTCCAGAGTGTTC-----TCTCNACAGCAGCAGCAACGACTACTCAGGA 763

QY 1803 TTTCCATAAAATGACGCTGAT 1824
DB 764 TTTCCATAAAATGACGCTGAT 785

RESULT 11
BB617224
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BB617224 692 bp mRNA linear EST 31-AUG-2001
BB617224 RIKEN full-length enriched, adult male testis Mus musculus
cDNA clone 4933431C08 5', mRNA sequence.

BB617224 GI:15396288

EST

BB617224.1

Mus musculus (house mouse)

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hananaki, T.,

Hara, A., Hiramoto, K., Hori, F., Ichi, Y., Ito, M., Kawai, J.,

Kanno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K.,

Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K.,

Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,

Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

RIKEN Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished (2001)

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,

Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,

Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A.

and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Kanno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,

Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A.,

Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K.,

Arakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.

Func. Genomics 2 pre, L72-L86 (2001)

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

e mouse tissues.

Location/Qualifiers

1. 692

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="4933431C08"

/sex="male"

/tissue_type="testis"

/lab_host="adult"

/dev_stage="SOLR"

/clone_lib="RIKEN full-length enriched, adult male testis"

/note="Site 1: XhoI; Site 2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia

Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in

RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGAGTCAAGAGCTCTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGAGCGCGCCGAATTAATTCGAGTAAATAATACCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites."

ORIGIN

Query Match	13.3%;	Score 424.4;	DB 2;	Length 692;
Best Local Similarity	79.7%;	Pred. No. 3.4e-111;		
Matches 553;	Conservative 0;	Mismatches 131;	Indels 10;	Gaps 4;

Qy	1004	AAACAAATATAGGCGAGTTGTGCGAGCGCTCTCGCAAGAGAAAGACAAGTGCAGAACTACCTCCC	1063
Db			
Qy	2	AAACAGTAAAGCAGTTGTCCAGGTCTCTCCCAAGACAAGAGAGCGCTTAGACATGCTA	61
Db			
Qy	1064	ACGCGAAGCGCCCTCCGTCGACCGGATCCCAAGTTTACCCAGTAGCAAGAGAGAAGGCCA	1123
Db			
Qy	62	ATAGTGAAGTGCCTTCTGGGATAGTGACCCCAAGTTGTCAGTAGCAAGAGAGGCCA	121
Db			
Qy	1124	CTCACTGCTCGAGTGGCGGCAAGCTTTTCAGAACTTACCACCACTGGTCTTTGCACCTCA	118
Db			
Qy	122	CGCACTGTTCTAGTGCAGCAAGAGCCTTTCAGGACATACCAACAGCTGCTCTGCACCTCA	181
Db			
Qy	1184	GGGTCCACAGAAGGACCGGAGGCGCGCGGAGTCCGCCACCATCTCTGTGTGACGGGA	1243
Db			
Qy	182	GGGTGCACAGGAAGCAGGAGGACTGATGCCCTGTCAACCACCATGGCTGTGGATGCA	241
Db			
Qy	1244	GGCAGCCGGGACGTGTTTCTCCTGACCTCGCGGCCCTCTGGATGAAAATGGAGCCGTGG	1303
Db			
Qy	242	GGCAGCTGGGACCTGTCTCCCGACACCTTCAGCACCATCTTGGAGACAGTGGGCGCGGG	301
Db			
Qy	1304	ATCAGAGGGAAGTGTTTCTGCAAGACGGAATCTGAGATGGGCTTCCGAGAGGAATTCATC	1363
Db			
Qy	302	ACCGA--GAGGGGGCTCTGAAGACGGGTCTGAGATGGACTCCCTGACGGGGTCCATT	358
Db			
Qy	1364	TGGATAAAATGATGATGGAGGAAAA-ATAAAACATCTTACATCTTCAAGAGAGTGTAGT	1422
Db			
Qy	359	TGGATAAAATGATGATGGAGGAAAGGGAAGCCCTCCCGTCTCTCGAGAGTGTAGT	418
Db			
Qy	1423	TATTGTGGAAAGTTTTTCCGTTCAAAATTAATTAACCTCAATATTCATCTCAGAACGCATACA	1482
Db			
Qy	419	TACTGTGGCAAGTTTTTCCGTTCAAACTATTATTAACCTCAATATTCATCTCAGAACGCATACA	478
Db			
Qy	1483	GGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCGCCAGAGACATCTCTG	1542
Db			
Qy	479	GGTGAAAAACCATACAAATGTGAATTTCTGTGAGTATGCCGAGCGCCAGAGACATCTCTG	538
Db			
Qy	1543	AGGTATCACTTGGAGAGACATCACAAAGGAAAAACAAG--CGATGTTGTGCTGCAAGTC	1599
Db			
Qy	539	AGGTACCACTTGGAGAGACATCACAAAGACAGCAGCGGTGGATGCTGCCGCTAGTCC	598
Db			
Qy	1600	AAGAACGATGTTAAAAATCAGGACACATGGAAGTCACTATTAA---CCGCTGCACAGTGCG	1656
Db			
Qy	599	AAAAGTGAAGCCGAGCCAGGCGCGCAGGATGCGCTACTAACGGCTGCTGCACAGTGCG	658
Db			
Qy	1657	CAAAACCAAAATTTGAAAGATTTTTTGATGGTG	1690
Db			
Qy	659	CAGACCAAAATTTAAAGAGATTTCTTTGAGGGTG	692
Db			

RESULT 12	BF237807	768 bp	mRNA	linear	EST 14-NOV-2000
LOCUS	601842326F1	NIH_MGC_46	Homo sapiens	cdna clone IMAGE:4079935 5',	
DEFINITION	mRNA sequence.				
ACCESSION	BF237807				
VERSION	BF237807.1				
KEYWORDS	GI:11151713				
	EST.				

```

SOURCE      Homo sapiens (human)
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE   1 (bases 1 to 768)
            NIH-MGC http://mgc.nci.nih.gov/
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC

CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: LLCW939 row: c column: 08
High quality sequence stop: 464.
Location/Qualifiers
    1..768
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:4079935"
    /tissue_type="leiomyosarcoma cell line"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NIH_MGC_46"
    /note="Organ: uterus; Vector: pOTB7; Site_1: XhoI; Site_2:
    EcoRI; cDNA made by oligo-dT priming. Directionally cloned
    into EcoRI/XhoI sites using the following 5' adaptor:
    GGCACGAG(G). Size-selected >500bp for average insert size
    1.8kb. Library constructed by Ling Hong in the laboratory
    of Gerald M. Rubin (University of California, Berkeley)
    using ZAP-cDNA synthesis kit (Stratagene) and Superscript
    II RT (Life Technologies). Note: this is a NIH_MGC
    Library."
FEATURES             source

```

ORIGIN

Query Match	13.1%;	Score 418.8;	DB 2;	Length 768;
Best Local Similarity	97.7%;	Pred. No. 1.5e-109;		
Matches 467;	Conservative	0;	Mismatches 7;	Indels 4;
Gaps	4;			
Qy	857	GCATCCCTCAGCTCGATTCGTTCCACCACCTTCACAGCTTGGCAGCTGGCTACCCAAAGGAA	916	
Db	1	GCATCCCTCAGCTCGATTCGTTCCACCACCTTCACAGCTTGGCAGCTGGCTACCCAAAGGAA	60	
Qy	917	AAGTTGCCAT-TTGCCAAAGAAGTGAAGGAATCGGGCAAGAAAGGGAGCACCGCAACACGAC	975	
Db	61	AAGTTGCCATATTGCCAAAGAAGTGAAGGAATCGGGCAAGAAAGGGAGCACCGCAACACGAC	120	
Qy	976	GATTCGAGTTCCAGAGAAGGCTTGGAGAACAAATAAGGGCAGTTGTGCAGGCTCTCG	1035	
Db	121	GATTCGAGTTCCAGAGAAGGCTTGGAGAACAAATAAGGGCAGTTGTGCAGGCTCTCTCG	180	
Qy	1036	CA-AGAGAAAGAGAGTGCAAAACACTCCCAACGCGCAAGCGCCCTCCGTGCAGCCGGATCC	1094	
Db	181	CATAGAAAGAGAGTGCAAAACACTCCCAACGCGCAAGCGCCCTCCGTGCAGCCGGATCC	240	
Qy	1095	CAAG-TTACCCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGGCGCAAAAGCTTTCA	1153	
Db	241	CAAGATTACCAGTAGCAAGGAGAAGCCCACTCACTGCTCCGAGTGGCGCAAAAGCTTTCA	300	
Qy	1154	GAACCTTACACACGAGCTGTTTGCATCCAGGGTCCACAGAGAGGACCCGAGAGGCGCGCG	1213	
Db	301	GAACCTTACACACGAGCTGTTTGCATCCAGGGTCCACAGAGAGGACCCGAGAGGCGCGCG	360	
Qy	1214	CGAGTCGCGCCACCATGTCGTGGACGGGAGGCAGCCGGGGAAGTGTTCCTCGACCTCG	1273	
Db	361	CGNGTGCGCCCAACATGTCGTGGACGGGAGGCAGCCGGGGAAGTGTTCCTCGACCTCG	420	
Qy	1274	CCGCCCCCTCTGGATGA-AAATGGAGCGCTGGATCGAGGGGAAGGTGTTCTGAAGACG	1330	

```

Db      421  CCGCCCCCTCTGGATGATACATCGAGCCGCTGGATCGAGGGAAGGCTGGTCTGAATAAG 478

RESULT 13
W05407
LOCUS   za83c12.r1 Soares fetal lung NBHL19W Homo sapiens cDNA clone
DEFINITION
IMAGE:299158 5' similar to SW:2F26_MOUSE P10076 ZINC FINGER PROTEIN
ZFP-26 ;, mRNA sequence.
ACCESSION
W05407.1 GI:1278138
VERSION 1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 473)
AUTHORS Hillier,L., Clark,N., Dubucque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: ETPRimer
High quality sequence stop: 406.
FEATURES
Location/Qualifiers
1..473
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1244082"
/db_xref="taxon:9606"
/clone="IMAGE:299158"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_lung_NBHL19W"
/note="Organ: lung; Vector: pT73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCAATTTTTTTTTTTT-3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by Bento
Soares and M. Fatima Ronaldo. This library was constructed
from the same fetus as the fetal heart library, Soares
fetal heart NBHL19W."

ORIGIN
Query Match 12.4%; Score 395.2; DB 7; Length 473;
Best Local Similarity 92.9%; Pred. No. 9e-103;
Matches 446; Conservative 0; Mismatches 27; Indels 7; Gaps 3;

Qy      1106 GTAGCAAGAGAGCCCACTACTCTCGAGTCGCGCAAAAGCTTTCAGAACCTACCACC 1165
Db      1 GTAGCAAGAGAGAGCCCACTACTCTCGAGTCGCGCAAAAGCTTTCAGAACCTACCACC 60

Qy      1166 AGCTGGTCTTGACCTCCAGGTCCACAGAGGACCGGCGCGGAGTGCCTCCCA 1225
Db      61 AGCTGGTCTTGACCTCCAGGTCCACAGAGGACCGGA-GGCGGTGCGGAGTCG-CCA 118

Qy      1226 CCATGCTGTGACGCGGAGGACCGGAGCGTGTCTCTGACCTCCGCCGCCCTCTGG 1285
Db      119 CCATGCTGTGACGCGGAGGACCGGAGCGTGTCTCTGACCTCCGCCGCCCTCTGG 178

```

```

Qy      1286 ATGAAATGGAGCCGCTGGATCGAGGGGAAGGTGTTCTGAAGACGGATCTGAGGATGGC 1345
Db      179 ATGAAATGGAGCCGCTGGATCGAGGGGAAGGTGTTCTGAAGACGGATCTGAGGATGGC 238

Qy      1346 TTCCCGAAGGAATCCATCTCGATAAAAATGATGAGGAGGAAAAATAAAACATCTTACAT 1405
Db      239 TTCCCGATAGATCCATCTCGATAAAAATGATGAGGAGGAAAAATAAAACATCTTACAT 298

Qy      1406 CTTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAAATTTTACCTCAATATTC 1465
Db      299 CTTCAAGAGAGTGTAGTTATTGTGAAAGTTTTTCCGTTCAAAATTTTACCTCAATATTC 358

Qy      1466 ATCTCAGAACCCATACAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAG 1525
Db      359 ATCTCAGAACCCATACAGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAG 418

Qy      1526 CCCAGNAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAAGGAAAAACAAACCGATG 1585
Db      419 CCAGNAGACATCTCTGAGGTATCACTTGGAGAGACATCACAAAGGAAAAACAAACCGATG 473

RESULT 14
BG007233
LOCUS   IL5-GN0239-271100-281-b03 GN0239 Homo sapiens cDNA, mRNA sequence.
DEFINITION
ACCESSION BG007233
VERSION BG007233.1 GI:12451213
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 411)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W.Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?ti=IL5&t2=IL5-GN0239-
271100-281-b03&t3=2000-11-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 360.
FEATURES
Location/Qualifiers
1..411
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0239"
/note="Organ: placenta_normal; Vector: puc18; Site 1:
SmaI; Site 2: SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196.716 - Ludwig Institute for Cancer
Research) profiles into the puc 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

```

ORIGIN

```
Query Match      12.4%; Score 394; DB 4; Length 411;
Best Local Similarity 97.3%; Pred. No. 1.9e-102; Indels 0; Gaps 0;
Matches 400; Conservative 0; Mismatches 11;

Qy 2433 CACTTTAGCCCCAAGTAACCTGAAGTCCACAGACCACAGAGAAATGTGGGGTCCAAGG 2492
Db 1 CACGAGCGCTCCAGTAACCTGAAGTCCACAGACCACAGAGAAATGTGGAGATCCAAGG 60

Qy 2493 GCGCCGACACAGGCGAACGAGATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCGTCGACC 2552
Db 61 GCGCCGACACAGGCGAACGAGATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCGTCGACC 120

Qy 2553 GGATAAGACAAAAGAGACCCGAGACAAAATGAAACCTCTTCAGTAGCTCTTCTTCAGCC 2612
Db 121 GGATAAGACAAATAGACCCGAGACAAAATGAAACCTCTTCAGTAGCTCTTCTTCAGCC 180

Qy 2613 CACCCTCGGACGAGTAACATCAATGGTTTCATCGACTACCCCGCAAGAACGACAGCCC 2672
Db 181 CACCCTCGGACGAGTAACATCAATGGTTTCATCGACTACCCCGCAAGAACGACAGCCC 240

Qy 2673 GTGGCCACCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATACCTGAGCAGA 2732
Db 241 GTGGCCACCTCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATACCTGAGCAGA 300

Qy 2733 ATTTGGTCAGCCCTTCCAAAAGAGCTCAAGTCCAGCGTGGTTGCCCTTGACGTTGACCA 2792
Db 301 ATTTGGTCAGCCCTTCCAAAAGAGCTCAAGTCCAGCGTGGTTGCCCTTGACGTTGACCA 360

Qy 2793 GCGCGGGCCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGT 2843
Db 361 GCGCGGNGCCCAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGT 411
```

RESULT 15

```
BM834501
LOCUS K-EST0109502 S11SNU1 Homo sapiens cDNA clone S11SNU1-63-D04 5',
DEFINITION mRNA sequence.
ACCESSION BM834501
VERSION BM834501.1 GI:19190910
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
```

21C Frontier Korean EST Project 2001

TITLE

Unpublished (2002)

JOURNAL

Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

COMMENT

Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 63 row: D column: 04
High quality sequence stop: 640.

FEATURES

```
Location/Qualifiers
1..640
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S11SNU1-63-D04"
/sex="M"
/tissue_type="Stomach"
/cell_type="Lymphoblast-like"
/lab_host="SNU-1"
/lab_host="Top10P"
```

/clone_lib="S11SNU1"

/note="Organ: Stomach; Vector: pME18-FL3; Site 1: XhoI; Site 2: XhoI; The poly (A) + RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then decapped with tobacco acid pyrophosphatase (TAP). The decapped intact mRNA was ligated with DNA-RNA linker including SfiI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized with Superscript II using SfiI oligo-dT primer. After first strand synthesis, RNA was degraded by NaOH treatment and cDNA was amplified by PCR reaction. The PCR products were digested with SfiI and cloned into BrallI- digested pME18S-FL3 vector. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

```
Query Match      12.1%; Score 384; DB 4; Length 640;
Best Local Similarity 100.0%; Pred. No. 1.8e-99;
Matches 384; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 60
Db 257 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 316

Qy 61 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
Db 317 CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 376

Qy 121 AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 180
Db 377 AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 436

Qy 181 GGGTATATGCCCTTGGATTGCAATGTCGAGCCAGACCTTCACATTTTCAGAGACCTT 240
Db 437 GGGTATATGCCCTTGGATTGCAATGTCGAGCCAGACCTTCACATTTTCAGAGACCTT 496

Qy 241 AATAAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACGAGAGTTCCTTCGGGTT 300
Db 497 AATAAACATGTCTTAATGCAACACCGGCTACCTCTGTGAACGAGAGTTCCTTCGGGTT 556

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 557 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 616

Qy 361 AAGAAATTCAGAGGAAATGAATTT 384
Db 617 AAGAAATTCAGAGGAAATGAATTT 640
```

Search completed: July 3, 2005, 16:40:02

Job time : 10081.7 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 13000 Seconds

(without alignments)

11428.001 Million cell updates/sec

Title: US-08-731-499-12

Perfect score: 3066

Sequence: 1 GGAAACAGCTATGACCATGA.....GGTACCCAAATCGCCCTATA 3066

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pi.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3066	100.0	3066	6	BD085735 Genes fro
2	2813	91.7	2813	6	CQ776693 Sequence
3	2813	91.7	2813	6	AR371889 Sequence
4	2813	91.7	2813	9	AF041260 Homo sapi
5	2740.4	89.4	2808	6	CQ726468 Sequence
6	1016.4	33.2	163196	9	AC005220 Homo sapi
7	912.8	29.8	2916	10	AY219233
8	584.6	19.1	39507	9	AC004501
9	583	19.0	181313	2	AC021963
10	470.6	15.3	1885	10	AB029897
11	470.4	15.3	488	6	AX150111
12	303.6	9.9	105616	10	AL935134
13	303.6	9.9	168876	2	AC126535
14	303.6	9.9	232333	2	AC094145
15	303.6	9.9	235411	2	AC084066
16	265	8.6	265	6	CQ112058 Sequence
17	265	8.6	265	6	CQ150869 Sequence
18	265	8.6	265	6	CQ234209 Sequence
19	265	8.6	265	6	CQ309538 Sequence

20	265	8.6	265	6	CQ346160	Sequence	
21	264.2	8.6	724	10	BC048647	Mus muscu	
22	260.4	8.5	495	6	CQ099004	Sequence	
23	260.4	8.5	495	6	CQ137928	Sequence	
24	260.4	8.5	495	6	CQ221300	Sequence	
25	260.4	8.5	495	6	CQ297119	Sequence	
26	260.4	8.5	495	6	CQ333492	Sequence	
27	247	8.1	349	6	BD060529	Secreted	
C	28	244.4	8.0	106018	9	HS864118	Human DNA
	29	241.2	7.9	228652	9	CNS01DWD	Human chr
	30	240.8	7.9	180914	9	CNS01DS1	Human chr
C	31	240.8	7.9	181183	2	AL391866	Homo sapi
	32	240.8	7.9	214527	9	CNS0000J	Human chr
C	33	240.6	7.8	686	11	AF003745	Human chr
C	34	238.4	7.8	138573	9	AL954211	Pan trogl
35	238.2	7.8	105208	9	AL355515	Human DNA	
36	238	7.8	297	6	AX898536	Sequence	
37	238	7.8	297	6	BD034069	Sequence	
38	237.6	7.7	172510	9	AL512424	Human DNA	
39	237.4	7.7	159667	9	AC078953	Homo sapi	
C	40	237.2	7.7	173804	9	HS1030M6	Human DNA
41	236.8	7.7	144189	9	HS417M14	Human DNA	
C	42	236.8	7.7	152322	2	AL357077	Homo sapi
C	43	236.8	7.7	196662	9	AP001107	Homo sapi
C	44	236.6	7.7	149253	2	AC012245	Homo sapi
45	236.6	7.7	176132	2	AC061976	Homo sapi	

ALIGNMENTS

RESULT 1
BD085735
LOCUS
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085735
VERSION BD085735.1 GI:22631345
KEYWORDS JP 2001524802-A/11.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 3066)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 11 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
COMMENT OS Artificial Sequence
PN JP 2001524802-A/11
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12Q1/68, A61K48/00
CC Description of Artificial Sequence:lb1
FT Key Location/Qualifiers
FT source 1..3066
FT /organism='Artificial Sequence'.
FEATURES
source Location/Qualifiers
1..3066
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'

ORIGIN
Query Match 100.0%; Score 3066; DB 6; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GGAAACAGCTATGACCATGATTACGCCAAGCTCGAATTAACTCACTTAAGGGAACAA 60

Db 1 GGAACAGCTATGACCATGATTAGCCCAAGCTCGAAATTAACCCCTCACTAAAGGGAACAA 60
Qy 61 AAGCTGGAGCTCCACCGGCTGGCGGCTCTAGAACTAGTGTGATCCCCGGGCTGCAG 120
Db 61 AAGCTGGAGCTCCACCGGCTGGCGGCTCTAGAACTAGTGTGATCCCCGGGCTGCAG 120
Qy 121 GAAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGACAGCAGCACTGGAGACCCAG 180
Db 121 GAAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGACAGCAGCACTGGAGACCCAG 180
Qy 181 GACCTGTGACAGAGCACTCCGGGTGACACGAGGGAAGTCCCAAGAGTGAAGACCAAGAGAT 300
Db 181 GACCTGTGACAGAGCACTCCGGGTGACACGAGGGAAGTCCCAAGAGTGAAGACCAAGAGAT 300
Qy 241 TCACAGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTGAAGACCAAGAGAT 300
Db 241 TCACAGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTGAAGACCAAGAGAT 300
Qy 301 GAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCAGTGGTG 360
Db 301 GAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCAGTGGTG 360
Qy 361 GTGTGACCCACACAGTTTACGACTTAGAGGAACTCGACTTGGGAATAAGTGTCAAGACG 420
Db 361 GTGTGACCCACACAGTTTACGACTTAGAGGAACTCGACTTGGGAATAAGTGTCAAGACG 420
Qy 421 GATAATGTGGCACTTCTTCCCGGACAGCAGGATAGTGTGGGTGCGGATGCCAAC 480
Db 421 GATAATGTGGCACTTCTTCCCGGACAGCAGGATAGTGTGGGTGCGGATGCCAAC 480
Qy 481 GGAAGAACTTTGGGAAGAGCCAAACCCGAGCACCAGCTGTCTAAATCTCGTTTTTTC 540
Db 481 GGAAGAACTTTGGGAAGAGCCAAACCCGAGCACCAGCTGTCTAAATCTCGTTTTTTC 540
Qy 541 TTGATGTCTCTCGGCTGTACAGGACGTACCGGAGCAAGCGCAGATTCATCCCTT 600
Db 541 TTGATGTCTCTCGGCTGTACAGGACGTACCGGAGCAAGCGCAGATTCATCCCTT 600
Qy 601 GGATCAGTGAAGCTTGATGTGAGTCCCAATTAAGCTCCAGGCAACAGACCCCAAGTGAG 660
Db 601 GGATCAGTGAAGCTTGATGTGAGTCCCAATTAAGCTCCAGGCAACAGACCCCAAGTGAG 660
Qy 661 AGCTGGACACTTCGCTGGAGCTGGACCGGGGAGGACACAGATATAAACCCCAAGGGCAC 720
Db 661 AGCTGGACACTTCGCTGGAGCTGGACCGGGGAGGACACAGATATAAACCCCAAGGGCAC 720
Qy 721 GCGCGGCCCCAAGACAAGGTCTCTCTGCGGCAAGGATCCACGCTTCTCCCACTGAG 780
Db 721 GCGCGGCCCCAAGACAAGGTCTCTCTGCGGCAAGGATCCACGCTTCTCCCACTGAG 780
Qy 781 ACAGGGGAGCAGGAGAGAGCTCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAAA 840
Db 781 ACAGGGGAGCAGGAGAGAGAGCTCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAAA 840
Qy 841 TTCTTTCAAGCTTGGACAGGAAAGGTGCCAGGTGACAGCAACAGGAAGCCCAAG 900
Db 841 TTCTTTCAAGCTTGGACAGGAAAGGTGCCAGGTGACAGCAACAGGAAGCCCAAG 900
Qy 901 AGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGCTTATCAGGSCAGTCCGATGAT 960
Db 901 AGGGCAGAGCATCAAGACAAGGTGGATGAGTTCCTGCTTATCAGGSCAGTCCGATGAT 960
Qy 961 GTCCCTGCGGAGGACATAGTTGACCGGACAGGAAAGAGGACAGCAAGTTCGAACT 1020
Db 961 GTCCCTGCGGAGGACATAGTTGACCGGACAGGAAAGAGGACAGCAAGTTCGAACT 1020
Qy 1021 GCGGATTCCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGAGATTCGCCAG 1080
Db 1021 GCGGATTCCTCTGCTCCCTGGGACCCAGAGGACTGGAGACTGCAAGAGGAGATTCGCCAG 1080
Qy 1081 GCAGCAGCTATAGCAGAGAAATAATTCATGATTTCTTTAAACCTCTGGTTTCA 1140

Db 1081 GCAGCAGCTATAGCAGAGAAATAATTCATGATTTCTTTAAACCTCTGGTTTCA 1140
Qy 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAAGTCAACCAC 1200
Db 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCAGAGACACCGGTGCTGAAAAAGTCAACCAC 1200
Qy 1201 ACTTACAGCTGAACTTAAAGTCCAGCAAAAGTCTTTTATCATCTCCAGGAGACCAAGGGCT 1260
Db 1201 ACTTACAGCTGAACTTAAAGTCCAGCAAAAGTCTTTTATCATCTCCAGGAGACCAAGGGCT 1260
Qy 1261 GGCAGAAATTCAAAGAGTCAACCCATCGGGGACACACAGTCCGTTGAAACCCCTGAA 1320
Db 1261 GGCAGAAATTCAAAGAGTCAACCCATCGGGGACACACAGTCCGTTGAAACCCCTGAA 1320
Qy 1321 CCTGCGAAGAGCAGCAACAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db 1321 CCTGCGAAGAGCAGCAACAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Qy 1381 TTTTGGAAAAAGTCAAGTAAAGAGGACTCAGTCCCAAGGTGCGGAGGAGAAATGTGGTG 1440
Db 1381 TTTTGGAAAAAGTCAAGTAAAGAGGACTCAGTCCCAAGGTGCGGAGGAGAAATGTGGTG 1440
Qy 1441 TGTGAGTCAACAGTAGAGATTATAAGTCCAAGGAAGTAGAACTCAGCTTTACAACAGTG 1500
Db 1441 TGTGAGTCAACAGTAGAGATTATAAGTCCAAGGAAGTAGAACTCAGCTTTACAACAGTG 1500
Qy 1501 GACCTCAACGAAGAGAGATGCTGACCTGAAACCCACAGAGGCAAACTCAAAAGAGAGAA 1560
Db 1501 GACCTCAACGAAGAGAGATGCTGACCTGAAACCCACAGAGGCAAACTCAAAAGAGAGAA 1560
Qy 1561 AGCAAAACAAAGAACTCTCTGATGGCTTTCTCAGACAAATGTCAAGTGAAGGGATGGA 1620
Db 1561 AGCAAAACAAAGAACTCTCTGATGGCTTTCTCAGACAAATGTCAAGTGAAGGGATGGA 1620
Qy 1621 GGGATCAACCACTCAGAGAAATAAATGGGAAGACTCCAGCTCCCAACATCAGACTCC 1680
Db 1621 GGGATCAACCACTCAGAGAAATAAATGGGAAGACTCCAGCTCCCAACATCAGACTCC 1680
Qy 1681 ACAGAAAGACTATCACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAA 1740
Db 1681 ACAGAAAGACTATCACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAA 1740
Qy 1741 GAGGCTCTCTCGAAGGACAAAGAGTCAAGAGTCCAGAGCCAGATGAACAGAGAGAGCAACAG 1800
Db 1741 GAGGCTCTCTCGAAGGACAAAGAGTCAAGAGTCCAGAGCCAGATGAACAGAGAGAGCAACAG 1800
Qy 1801 CAGGAAGCAAGAGAACCCAGCCAGTGCACAGAGCAGGCGGAGTGAACCAAGTCACTG 1860
Db 1801 CAGGAAGCAAGAGAACCCAGCCAGTGCACAGAGCAGGCGGAGTGAACCAAGTCACTG 1860
Qy 1861 CAGAAATGGGACAAAGCTCCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGCTTC 1920
Db 1861 CAGAAATGGGACAAAGCTCCAAAGAGACCTGAGAAAGCGGACAGTCCCTTGGGGCTTC 1920
Qy 1921 TTTAAAGCCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAACAGAGCCAGATATCC 1980
Db 1921 TTTAAAGCCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAACAGAGCCAGATATCC 1980
Qy 1981 ATCGGACAGTGGCAAAACCCAGTAAACAAATCAGCAGCGTTCCCAAGGTTCTCCTG 2040
Db 1981 ATCGGACAGTGGCAAAACCCAGTAAACAAATCAGCAGCGTTCCCAAGGTTCTCCTG 2040
Qy 2041 CCACAAAGATGTGTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Db 2041 CCACAAAGATGTGTTCTCTTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2100
Qy 2101 TTCTGATGGCCAGCAAAATGAAATTTGCTCTAGAAATTTAAGCCCGAGCTGTTGTATATGA 2160
Db 2101 TTCTGATGGCCAGCAAAATGAAATTTGCTCTAGAAATTTAAGCCCGAGCTGTTGTATATGA 2160
Qy 2161 GGTGTATTTATTTAGTCTCTGCTCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTA 2220
Db 2161 GGTGTATTTATTTAGTCTCTGCTCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTA 2220

Qy 976 GACATAGTTGACGCAAGGAAAAAGAGCAAGAACTTGAACCTGGGATGCTCTGTC 1035
Db |||||
Qy 841 GACATAGTTGACGCAAGGAAAAAGAGCAAGAACTTGGAACTGGGATGCTCTGTC 900
Db |||||
Qy 1036 CCTGGGACCCAGAAAGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 1095
Db |||||
Qy 901 CCTGGGACCCAGAAAGACTGGAGACTGCAAGAGACGATTCGCCAGGACGAGCTATAGCA 960
Db |||||
Qy 1096 GAGAAATAAATTCATCATGAGTTCTTTTAAACTCTGGTTTTCACCTTAAACAAAGCTGAA 1155
Db |||||
Qy 961 GAGAAATAAATTCATCATGAGTTCTTTTAAACTCTGGTTTTCACCTTAAACAAAGCTGAA 1020
Db |||||
Qy 1156 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCCACCTTCAGCTGACCTT 1215
Db |||||
Qy 1021 ACAAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCCACCTTCAGCTGACCTT 1080
Db |||||
Qy 1216 AAGTCAGAACAAAGCCNACTTTACATCCAGAGAGACCCAGGGCTGGCAAGATTCNAAA 1275
Db |||||
Qy 1081 AAGTCAGAACAAAGCCNACTTTACATCCAGAGAGACCCAGGGCTGGCAAGATTCNAAA 1140
Db |||||
Qy 1276 GGATGCAACCCATCGGGGCACACACAGTCCGCTGCAACCCCTGAACTTGGCAAGGAGGC 1335
Db |||||
Qy 1141 GGATGCAACCCATCGGGGCACACACAGTCCGCTGCAACCCCTGAACTTGGCAAGGAGGC 1200
Db |||||
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db |||||
Qy 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
Db |||||
Qy 1396 GTTAAAGAGGACTGATGCTCCACAGGTGCGAGGAGAAATGTTGTGAGTCAACAGTA 1455
Db |||||
Qy 1261 GTTAAAGAGGACTGATGCTCCACAGGTGCGAGGAGAAATGTTGTGAGTCAACAGTA 1320
Db |||||
Qy 1456 GAGATTATAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGTGGACTCAACGAAGGA 1515
Db |||||
Qy 1321 GAGATTATAAGTCCAAAGGAGTAGAATCAGCTTTACAAACAGTGGACTCAACGAAGGA 1380
Db |||||
Qy 1516 GATGTGCACCTGAACCCACAGAGAGCAAACTCAAAAGAGAGAAAGCAAAACCAAGAAC 1575
Db |||||
Qy 1381 GATGTGCACCTGAACCCACAGAGAGCAAACTCAAAAGAGAGAAAGCAAAACCAAGAAC 1440
Db |||||
Qy 1576 TCTCTGATGGGTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCACCCACTCA 1635
Db |||||
Qy 1441 TCTCTGATGGGTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCACCCACTCA 1500
Db |||||
Qy 1636 GAAGAAATAAATGGGAAAGACTCCAGCTGCAACATCAGACTCCACAGAAAGAGACTATC 1695
Db |||||
Qy 1501 GAAGAAATAAATGGGAAAGACTCCAGCTGCAACATCAGACTCCACAGAAAGAGACTATC 1560
Db |||||
Qy 1696 ACACGCCAGAGCTTGAAACCAACAGGAGCACACAGAGGGGTAAAGAGGGCTCTCTCGAAG 1755
Db |||||
Qy 1561 ACACGCCAGAGCTTGAAACCAACAGGAGCACACAGAGGGGTAAAGAGGGCTCTCTCGAAG 1620
Db |||||
Qy 1756 GACAAGAGTACAGAGCCGAGATGAACAGAGAGAGCAACAGAGAGAGCCCAAGAA 1815
Db |||||
Qy 1621 GACAAGAGTACAGAGCCGAGATGAACAGAGAGAGCAACAGAGAGAGCCCAAGAA 1680
Db |||||
Qy 1816 CCAGCCAGTGCACAGAGCAGCCAGTGGACAGCAACTCACTCGCAATGGGGACAAG 1875
Db |||||
Qy 1681 CCAGCCAGTGCACAGAGCAGCCAGTGGACAGCAACTCACTCGCAATGGGGACAAG 1740
Db |||||
Qy 1876 CTCAAAAAGAGACCTTGAGAGCGGCGAGAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1935
Db |||||
Qy 1741 CTCAAAAAGAGACCTTGAGAGCGGCGAGAGTCCCTTGGGGGCTCTTTAAAGGCTGGGA 1800
Db |||||
Qy 1936 CCAAGCGGATGTTGGATGCTCAAGTGCAACAGAGCCAGATATCATCGACAGTTCGCTGGC 1995
Db |||||
Qy 1801 CCAAGCGGATGTTGGATGCTCAAGTGCAACAGAGCCAGATATCATCGACAGTTCGCTGGC 1860
Db |||||
Qy 1996 AAACCCAGTAAACAAATCAGACCGGTTCCACAGGTTCTCTGCCACCAAGATGTGT 2055
Db |||||
Qy 1861 AAACCCAGTAAACAAATCAGACCGGTTCCACAGGTTCTCTGCCACCAAGATGTGT 1920
Db |||||
Qy 2056 CTCCTTACTCCATCTCTCTCCCAACACAGCTCCATGTATATATTTCTTGATGGCCAGCA 2115

Db 1921 CTCCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATATTTCTTGATGCCAGCA 1980
Qy |||||
Qy 2116 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTCTGTATATTTAGGTGTATTTTACG 2175
Db |||||
Qy 1981 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTCTGTATATTTAGGTGTATTTTACG 2040
Db |||||
Qy 2176 TCTCTGCTCAGCTCTTTTGGCAAAATAACAGATAAGATGGTTTTAGCAGGTCACCTAGTT 2235
Db |||||
Qy 2041 TCTCTGCTCAGCTCTTTTGGCAAAATAACAGATAAGATGGTTTTAGCAGGTCACCTAGTT 2100
Db |||||
Qy 2236 GGCTCAGAAGAGTCGATGATCACAAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2295
Db |||||
Qy 2101 GGCTCAGAAGAGTCGATGATCACAAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2160
Db |||||
Qy 2296 GTTAAGTGTATGAAATTCGAGTGTGGCGGCTGTGGCTCTCGCTGTAACTCTCAGC 2355
Db |||||
Qy 2161 GTTAAGTGTATGAAATTCGAGTGTGGCGGCTGTGGCTCTCGCTGTAACTCTCAGC 2220
Db |||||
Qy 2356 ACTTTGGGAGGCCGAGGACGAGTGGATCACCTGAGGTCAGGAGTTCAAGACTAGCTTGCC 2415
Db |||||
Qy 2221 ACTTTGGGAGGCCGAGGACGAGTGGATCACCTGAGGTCAGGAGTTCAAGACTAGCTTGCC 2280
Db |||||
Qy 2416 AACATCATGAAACCCCGTCTCTACTTAAATAACAAAAATAGCCAGGATGTTGGCACAC 2475
Db |||||
Qy 2281 AACATCATGAAACCCCGTCTCTACTTAAATAACAAAAATAGCCAGGATGTTGGCACAC 2340
Db |||||
Qy 2476 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2535
Db |||||
Qy 2341 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAGGTG 2400
Db |||||
Qy 2536 GAGGTTCAGTGTAGCCGAAAGTTGCACATTGCACTTCCACCTCGGCGCAGACAGCAAGATT 2595
Db |||||
Qy 2401 GAGGTTCAGTGTAGCCGAAAGTTGCACATTGCACTTCCACCTCGGCGCAGACAGCAAGATT 2460
Db |||||
Qy 2596 CTATCAAAAAAAGAGGAGTGGCAAGTAAAGTATAGAGAGAAATGCTGTAGNAGGAA 2655
Db |||||
Qy 2461 CTATCAAAAAAAGAGGAGTGGCAAGTAAAGTATAGAGAGAAATGCTGTAGNAGGAA 2520
Db |||||
Qy 2656 TTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTTGAAGAGGAGAGCAAAATCC 2715
Db |||||
Qy 2521 TTAAGCGTTGTAGTAAACCGTCTCATCTCTAAGCTTTGAAGAGGAGAGCAAAATCC 2580
Db |||||
Qy 2716 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGCGCTGTGTGGTGGTTGCCAAT 2775
Db |||||
Qy 2581 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGCGCTGTGTGGTGGTTGCCAAT 2640
Db |||||
Qy 2776 TTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATTAAGCGTTGTTTTCAAT 2835
Db |||||
Qy 2641 TTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATTAAGCGTTGTTTTCAAT 2700
Db |||||
Qy 2836 AGGTTCCTTTGTAAGTTATTGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAATGAT 2895
Db |||||
Qy 2701 AGGTTCCTTTGTAAGTTATTGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAATGAT 2760
Db |||||
Qy 2896 TTGGGAAACAAATTCCTTTTTCAGGCTCAGTCACAAACGCAAGATTAACAATT 2948
Db |||||
Qy 2761 TTGGGAAACAAATTCCTTTTTCAGGCTCAGTCACAAACGCAAGATTAACAATT 2813
Db |||||

RESULT 3

AR371889 LOCUS AR371889 2813 bp DNA linear PAT 12-SEP-2003
DEFINITION Sequence 3 from patent US 6395544.
ACCESSION AR371889
VERSION AR371889.1 GI:34608999
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2813)
AUTHORS Cowser, L.M. and Freier, S.M.
TITLE Antisense modulation of BCAS1 expression

Db 361 AAAGAGCCCAACCCGAGGACCAGCTGCTAAATCTCGTTTTTTCTTGATGCTCTCTCGG 420
Qy 556 CCTGTACAGAGAGCTACCGGAGACCAAGCCGAGATTCATCCCTTGGATFCAGTGAAGCTT 615
Db 421 CCTGTACAGAGAGCTACCGGAGACCAAGCCGAGATTCATCCCTTGGATFCAGTGAAGCTT 480
Qy 616 GATGTCAAGCTCCAAATAAAGCTCCAGCGAACAAAGACCCCAAGTGAGAGCTGGACACTTCGG 675
Db 481 GATGTCAAGCTCCAAATAAAGCTCCAGCGAACAAAGACCCCAAGTGAGAGCTGGACACTTCGG 540
Qy 676 GTGCGAGCTGGACCGGGGAGGACACAGATAAACCCCAAGGAGCAGCCCGCCGCCCCAAGAC 735
Db 541 GTGCGAGCTGGACCGGGGAGGACACAGATAAACCCCAAGGAGCAGCCCGCCGCCCCAAGAC 600
Qy 736 AAGTCTCTCTGCGCCGAGGATCCAGAGCTTCTCCACTGTGAGACAGGGGAGCAGGA 795
Db 601 AAGTCTCTCTGCGCCGAGGATCCAGAGCTTCTCCACTGTGAGACAGGGGAGCAGGA 660
Qy 796 GGAGAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATTTCTCAAGCTGGAC 855
Db 661 GGAGAGCTCCCTCCAAAGCCCAAGACTCCAGCTTTTGTGACAAATTTCTCAAGCTGGAC 720
Qy 856 AAGGACAGGAAAAGGTCCAGGTGACAGCCCAAGAGAAAGCCAGAGGCGCAGAGCATCAA 915
Db 721 AAGGACAGGAAAAGGTCCAGGTGACAGCCCAAGAGAAAGCCAGAGGCGCAGAGCATCAA 780
Qy 916 GACAGGTGGATGAGGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTCTGACGGGAAG 975
Db 781 GACAGGTGGATGAGGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTCTGACGGGAAG 840
Qy 976 GACATAGTTGACGGCAAGGAAAAGAGGACAGAGCAAGAACTTGGAACTGGCGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGGCAAGGAAAAGAGGACAGAGCAAGAACTTGGAACTGGCGGATGCTCTGTC 900
Qy 1036 CTGCGGACCCAGAGGACTGGAGACTGCAAGAGACAGATTCCTCCAGGCGAGCAGCTATAGCA 1095
Db 901 CTGCGGACCCAGAGGACTGGAGACTGCAAGAGACAGATTCCTCCAGGCGAGCAGCTATAGCA 960
Qy 1096 GAGATAATAATTCATCATGAGTTCTTTAAACTCTGTTTACCTTAAACAAAGCTGAA 1155
Db 961 GAGATAATAATTCATCATGAGTTCTTTAAACTCTGTTTACCTTAAACAAAGCTGAA 1020
Qy 1156 ACNAAAAGGACCCAGAGACACGGTGTGAAAAGTCAACCCACCTTACGCTGACCTT 1215
Db 1021 ACNAAAAGGACCCAGAGACACGGTGTGAAAAGTCAACCCACCTTACGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAACTTCGGAAGGAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCGTCGACAAACCCCTGAACTTCGGAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGGCTGGGAGGAGATGTTGTTGAGTCAACCACTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGGCTGGGAGGAGATGTTGTTGAGTCAACCACTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAAGTAGAATCAGCTTTACAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAAGTAGAATCAGCTTTACAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATGTGCACTGAAACCCACAGAGAGGAAAACCTCAAAAGAGAGAAAGCAAAACCAAGAAC 1575
Db 1381 GATGTGCACTGAAACCCACAGAGAGGAAAACCTCAAAAGAGAGAAAGCAAAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCAACCACTCA 1635
Db 1441 TCTCTGATGGCTTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATCAACCACTCA 1500

Qy 1636 GAAGAAATAATGGGAAAGACTCCAGTGTGCCAAACATCAGATCTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAATGGGAAAGACTCCAGTGTGCCAAACATCAGATCTCCACAGAAAAGACTATC 1560
Qy 1696 ACACGCCAGAGCTGTGAACCAACAGGAGCACCACAGAGGGTAAAGAGGGCTCCTCGAAG 1755
Db 1561 ACACGCCAGAGCTGTGAACCAACAGGAGCACCACAGAGGGTAAAGAGGGCTCCTCGAAG 1620
Qy 1756 GACAAGAGTCCAGCAGCGGAGATGAACAAAGCAGAGAGCAACAAAGCAGAGAAAGCAAGAA 1815
Db 1621 GACAAGAGTCCAGCAGCGGAGATGAACAAAGCAGAGAGCAACAAAGCAGAGAAAGCAAGAA 1680
Qy 1816 CCAGCCAGTCCACAGAGCAGGCCACGGTGACACGAACTCCTCAGAGAAATGGGGACAAAG 1875
Db 1681 CCAGCCAGTCCACAGAGCAGGCCACGGTGACACGAACTCCTCAGAGAAATGGGGACAAAG 1740
Qy 1876 CTCCAAAGAGACCTGTGAAGCGGACGAGTCCCTTGGGGCTCTTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGTGAAGCGGACGAGTCCCTTGGGGCTCTTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTCGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTCGC 1860
Qy 1996 AAAACCAAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAGATGTGT 2055
Db 1861 AAAACCAAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAGATGTGT 1920
Qy 2056 CTCTTACTCCTCTCTCCCAACAGCTCCATGTATATATTTCTTGTGATGCCAGCA 2115
Db 1921 CTCTTACTCCTCTCTCCCAACAGCTCCATGTATATATTTCTTGTGATGCCAGCA 1980
Qy 2116 AATGAAATTCCTCTAGAAATTAAGCCGAGCTGTGTATATTGAGGTGTATTTTACG 2175
Db 1981 AATGAAATTCCTCTAGAAATTAAGCCGAGCTGTGTATATTGAGGTGTATTTTACG 2040
Qy 2176 TCTCTGTCTCAGTCTTTCTTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGTCTCAGTCTTTCTTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAGAGAGTGCATGATCACCAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTCAGAGAGTGCATGATCACCAGCAGAAAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAAGTGAATAATGGCAGTGTGCGCGGCTGTGGCTCTCGCTGTAACTCAGC 2355
Db 2161 GTTAAAGTGAATAATGGCAGTGTGCGCGGCTGTGGCTCTCGCTGTAACTCAGC 2220
Qy 2356 ACTTTGGAGGCGCAGGACAGTGGATCACTTGAGGTCAAGAGTTCAAGACTAGCTGCGC 2415
Db 2221 ACTTTGGAGGCGCAGGACAGTGGATCACTTGAGGTCAAGAGTTCAAGACTAGCTGCGC 2280
Qy 2416 AACATCATGAACCCCTCTCTACTATAAATAACAAATAATAGCCAGGATGGTGGCACAC 2475
Db 2281 AACATCATGAACCCCTCTCTACTATAAATAACAAATAATAGCCAGGATGGTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAGACCGCTTGTACCCAGAGGTG 2535
Db 2341 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAGACCGCTTGTACCCAGAGGTG 2400
Qy 2536 GAGTTCAGTGGAGCCGAAGTTGCACCTTGTCACTCCACCTCGGGCGCAGAGCAAGATT 2595
Db 2401 GAGTTCAGTGGAGCCGAAGTTGCACCTTGTCACTCCACCTCGGGCGCAGAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGGAGTGAAGTTATAGAGAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGAAGTTATAGAGAGAAATGCTGTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACCGGTGCTCATCTCTAAGCTTGAAGAGGAGGAGCAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACCGGTGCTCATCTCTAAGCTTGAAGAGGAGGAGCAAAATCC 2580

```
Qy 2716 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGGGCTGTGTGTGGTGGTTGCCAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGAGGAGGAAACCCGGGCTGTGTGTGGTGGTTGCCAAT 2640
Qy 2776 TTCTAGAACCGAATGTGTGGGTATAGAAAAGGAATGAATAGCGTGTGTTTTTCAAAT 2835
Db 2641 TTCTAGAACCGAATGTGTGGGTATAGAAAAGGAATGAATAGCGTGTGTTTTTCAAAT 2700
Qy 2836 AGGTTCCTTGTAAAGTTATTTGATGAGAGGAAAGATTGACTGGGAGGGCTTAAAAATGAT 2895
Db 2701 AGGTTCCTTGTAAAGTTATTTGATGAGAGGAAAGATTGACTGGGAGGGCTTAAAAATGAT 2760
Qy 2896 TTGGGAAACAAATTCCTTTTGGAGCTCAGTGACAAACCGCAAGATTACAACTT 2948
Db 2761 TTGGGAAACAAATTCCTTTTGGAGCTCAGTGACAAACCGCAAGATTACAACTT 2813

RESULT 5
CQ726468          2808 bp      DNA      linear      PAT 03-FEB-2004
DEFINITION      Sequence 12402 from Patent WO02068579.
ACCESSION      CQ726468
VERSION        CQ726468.1  GI:42289602
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL      Patent: WO 02068579-A 12402 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. 2808
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match      89.4%; Score 2740.4; DB 6; Length 2808;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 2784; Conservative 0; Mismatches 21; Indels 11; Gaps 2;

Qy 136 CTCACCCAGCAGCAGGCACTGGGAGCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 195
Db 1 CTCCACCCAGCAGCAGGCACTGGGAGCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 60
Qy 196 CAGCTCCGGGTGACACGAGGGGACTGGAAGATACTCCACAGGGGTCTCAGCAGGAGCAATG 255
Db 61 CAGCTCCGGGTGACACGAGGGGACTGGAAGATACTCCACAGGGGTCTCAGCAGGAGCAATG 120
Qy 256 GGTAAACCAATAGTGTGTTCCCAAGAGTGTGAAGACCAAGAGAATGAACCAAGAGCAGAG 315
Db 121 GGTAAACCAATAGTGTGTTCCCAAGAGTGTGAAGACCAAGAGAATGAACCAAGAGCAGAG 180
Qy 316 ACTTACCAAGGACAACCGCTCTGCTCTGAACCGGGTTCAGTGGTGTGTCACCCACACA 375
Db 181 ACTTACCAAGGACAACCGCTCTGCTCTGAACCGGGTTCAGTGGTGTGTCACCCACACA 240
Qy 376 GTTCAGCACTTAGAGGAAAGTGCATTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAAGTGCATTGGGAATAAGTGTCAAGACGGATAAATGTGGCCACT 300
Qy 436 TCTTCCCGGAGACAAACCGAGATAAGTGTGTGGGATGTCGCAACCGAAAGAAATCTTGGG 495
Db 301 TCTTCCCGGAGACAAACCGAGATAAGTGTGTGGGATGTCGCAACCGAAAGAAATCTTGGG 360
Qy 496 AAAGAGGCCAAACCCGAGGCCAGCTGTAATCTCGTTTTTCTTGATGCTCTCTCGG 555
Db 361 AAAGAGGCCAAACCCGAGGCCAGCTGTAATCTCGTTTTTCTTGATGCTCTCTCGG 420
```

```
Qy 556 CCTGTACCGAGCGTATCCGGAGACCAAGCCGCGAGATTTCATCCCTTGGATCAGTGAAGCTT 615
Db 421 CCTGTACCGAGCGTATCCGGAGACCAAGCCGCGAGATTTCATCCCTTGGATCAGTGAAGCTT 480
Qy 616 GATGTCAAGTCCCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTAGAGCTGGACACTTCCG 675
Db 481 GATGTCAAGTCCCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTAGAGCTGGACACTTCCG 540
Qy 676 GTGCGAGCTCGACCGGGCAGGACACAGATAAAACCCAGGCGACGCCCGGCCCAAGAC 735
Db 541 GTGCGAGCTCGACCGGGCAGGACACAGATAAAACCCAGGCGACGCCCGGCCCAAGAC 600
Qy 736 AAGTCTCTCTCTGCGCCAGGGATCCCAAGCTTCTCCACCTCAGACAGGGGAGCAGGA 795
Db 601 AAGTCTCTCTCTGCGCCAGGGATCCCAAGCTTCTCCACCTCAGACAGGGGAGCAGGA 660
Qy 796 GGAGAAGCTCCCTCAAGCCCAAGGACTCCAGCTTTTTTTTTCACAAATTTCTTCAAGCTGGAC 855
Db 661 GGAGAAGCTCCCTCAAGCCCAAGGACTCCAGCTTTTTTTTTCACAAATTTCTTCAAGCTGGAC 720
Qy 856 AAGGACAGGAAAGAGTGGCCAGGTGACAGCCAAACAGGAAGCCAAAGAGGCGAGACATCAA 915
Db 721 AAGGACAGGAAAGAGTGGCCAGGTGACAGCCAAACAGGAAGCCAAAGAGGCGAGACATCAA 780
Qy 916 GACAAGTGTGATGAGGTTCCTTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCGAGGAAG 975
Db 781 GACAAGTGTGATGAGGTTCCTTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCGAGGAAG 840
Qy 976 GACATAGTTGACGGCAAGGAAAAAGAGGACAGAACTTGGAACTGCGGATGCTCTGTGTC 1035
Db 841 GACATAGTTGACGGCAAGGAAAAAGAGGACAGAACTTGGAACTGCGGATGCTCTGTGTC 900
Qy 1036 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGCGATTCCCAAGGCGAGCAGCTATAGCA 1095
Db 901 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGCGATTCCCAAGGCGAGCAGCTATAGCA 960
Qy 1096 GAGAATAATAATTCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAACAAAGCTGAA 1155
Db 961 GAGAATAATAATTCATCATGAGTTTCTTTAAACTCTGGTTTCACTTAACAAAGCTGAA 1020
Qy 1156 ACAAAGGACCCAGAGGACACCGGGTGTGAAAGTCAACCCACCTTCAAGCTGAGCTT 1215
Db 1021 ACAAAGGACCCAGAGGACACCGGGTGTGAAAGTCAACCCACCTTCAAGCTGAGCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCAAAGGGCTGGCAAGATTCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCAAAGGGCTGGCAAGATTCAAA 1140
Qy 1276 GGATGCAACCCATCGGGGCGACACACAGTCCGTGCAACCCCTGAACTGCGAAGGAAGGC 1335
Db 1141 GGATGCAACCCATCGGGGCGACACACAGTCCGTGCAACCCCTGAACTGCGAAGGAAGGC 1200
Qy 1336 ACCAAGAGAAATCAGAGACCCACCTCTCTGCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db 1201 ACCAAGAGAAATCAGAGACCCACCTCTCTGCTGGGCAAACTGTTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGAGGAGGAGTGTGTGTGAGTCAACCACTCA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGAGGAGGAGTGTGTGTGAGTCAACCACTCA 1312
Qy 1456 GAGATTATAAGTCCAAAGGAGTGAATCAGCCTTACAAACAGTGGACCTCAACGAAGGA 1515
Db 1313 GAGATTATAAGTCCAAAGGAGTGAATCAGCCTTACAAACAGTGGACCTCAACGAAGGA 1372
Qy 1516 GATGCTGACCTGAAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1575
Db 1373 GATGCTGACCTGAAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1432
Qy 1576 TCTCTGATGGGTTTCTCAGACAAATGTGTCAGTGAAGGGGATGAGGATCACCACCTCA 1635
Db 1433 TCTCTGATGGGTTTCTCAGACAAATGTGTCAGTGAAGGGGATGAGGATCACCACCTCA 1492
```

1636 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1695
1493 GAAGAAATAAATGGGAAGACTCCAGCTGCCAAAACATCAGACTCCACAGAAAAGACTATC 1552
1696 ACACCGCCAGAGCCTGAAACCAACAGAGAGCACCACAGAGGGGTAAAGAGGGCTTCCTCGAAG 1755
1553 ACACCGCCAGAGCCTGAAACCAACAGAGAGCACCACAGAGGGGTAAAGAGGGCTTCCTCGAAG 1612
1756 GACAGAGTTCAGCAGCCGAGATGAAACAGAGAGAGAGCAACAGAGGAAGCCAAAGAA 1815
1613 GACAGAGTTCAGCAGCCGAGATGAAACAGAGAGAGAGCAACAGAGGAAGCCAAAGAA 1672
1816 CCAGCCCACTGCACAGAGCAGCCAGCGTGGACACGAACTCACTCGAATGGGACCAAG 1875
1673 CAGCCCACTGCACAGAGCAGCCAGCGTGGACACGAACTCACTCGAATGGGACCAAG 1732
1876 CTCCTTACTTCCATCTCTCCCAACACAGCTCCATGTATATATCTTGGGGCTCTTTAAAGCCCTGGGA 1935
1733 CTCCTTACTTCCATCTCTCCCAACACAGCTCCATGTATATATCTTGGGGCTCTTTAAAGCCCTGGGA 1792
1936 CCAAGCCGATGTTGATGCTCAAGTGCAACAGACCCAGTATCCATCGACCACTGGC 1995
1793 CCAAGCCGATGTTGATGCTCAAGTGCAACAGACCCAGTATCCATCGACCACTGGC 1852
1996 AAACCCCAAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCTGCCACCAAGATGTGT 2055
1853 AAATCCAGTAAACAAATCAGCACGCTTCCACACAGGTTCTCTGCCACCAAGATGTGT 1912
2056 CTCCTTACTTCCATCTCTCCCAACACAGCTCCATGTATATATCTTGGGGCTCTTTAAAGCCCTGGGA 2115
1913 CTCCTTACTTCCATCTCTCCCAACACAGCTCCATGTATATATCTTGGGGCTCTTTAAAGCCCTGGGA 1972
2116 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTTATTTACG 2175
1973 AATGAAATTCCTGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTTATTTACG 2032
2176 TCTCTGCTCCAGTCTTTCTGCAATTAACAGTAAAGTGGTTTACGAGGTCACTAGTT 2235
2033 TCTCTGCTCCAGTCTTTCTGCAATTAACAGTAAAGTGGTTTACGAGGTCACTAGTT 2092
2236 GGGTCAGAGAGTCGATGATACCAAGCAGAGGAGGGAATAGAGGAATGTGTCGG 2295
2093 GGGTCAGAGAGTCGATGATACCAAGCAGAGGAGGGAATAGAGGAATGTGTCGG 2152
2296 GTTAAGTGAATAAATGACAGTGGTGGCGGCGTGGTGGCTCTCGCTGTAATCTCAGC 2355
2153 GTTAAGTGAATAAATGACAGTGGTGGCGGCGTGGTGGCTCTCGCTGTAATCTCAGC 2212
2356 ACTTTGGAGGCGGAGGAGGAGTGGATCACCTGAGGTCAGGATTCAGACTAGCCCTGGCC 2415
2213 ACTTTGGAGGCGGAGGAGGAGTGGATCACCTGAGGTCAGGATTCAGACTAGCCCTGGCC 2272
2416 AACATCATGAACCCCGTCTCTACTAAATAACAAAATTAAGCAGGATGGTGGCACAC 2475
2273 AACATCATGAACCCCGTCTCTACTAAATAACAAAATTAAGCAGGATGGTGGCACAC 2332
2476 ACCTGTAGTCCAGTACTCGGAGCCCAACAGCAGAACCGCTTTGACCCAGAGGTG 2535
2333 ACCTGTAGTCCAGTACTCGGAGCCCAACAGCAGAACCGCTTTGACCCAGAGGTG 2392
2536 GAGGTTCAGTGAAGCGAAGTTGCACCATTTGCACTCCACCTGGGCGCAGAGCAAGATT 2595
2393 GAGGTTCAGTGAAGCGAAGTTGCACCATTTGCACTCCACCTGGGCGCAGAGCAAGATT 2452
2596 CTATC- - -AAAAAAGAGGAGTGGCAGTAAAGTTATAGAGAGAAATGCTGTAGAAG 2652
2453 CTATCAAAAAAAGAGGAGTGGCAGTAAAGTTATAGAGAGAAATGCTGTAGAAG 2512
2653 GAATTAAGCGTTGTAGTAACCGCTGCTCATCTCTTAAGCTTTGAAGAAGGAGCAAAAA 2712
2513 GAATTAAGCGTTGTAGTAACCGCTGCTCATCTCTTAAGCTTTGAAGAAGGAGCAAAAA 2572
2713 TCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGTTGCC 2772

Db 2573 TCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGTTGCC 2632
Qy 2773 AATTTCTAGAACCGAATGTGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCA 2832
Db 2633 AATTTCTAGAACCGAATGTGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCA 2692
Qy 2833 AATAGGTCCTTGTAAAGTTATTTGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAAT 2892
Db 2693 AATAGGTCCTTGTAAAGTTATTTGATGAGAGGAAAAAGATTGACTGGGAGGCTTAAAAAT 2752
Qy 2893 GATTTGGGAAAAAATTCCTTTTGGAGCTCAGTGACACGCAACAGATTACAACTT 2948
Db 2753 GATTTGGGAAAAAATTCCTTTTGGAGCTCAGTGACACGCAACAGATTACAACTT 2808

RESULT 6
AC005220
LOCUS Homo sapiens chromosome 20, BAC clone 99 (LBNL H80), complete
DEFINITION sequence.
ACCESSION AC005220
VERSION AC005220.1 GI:3282159
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163196)
AUTHORS Gray, J.W., Collins, C., Kimmerly, W., Bondoc, M., Cheng, J.,
Connolly, K.S., Gunning, K.M., Kadner, K., Miguel, T., Miller, C.,
Pitluck, S., Pollard, M., Rojeski, H., Subramanian, S. and Martin, C.H.
TITLE Sequencing of human chromosome 20
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163196)
AUTHORS Ricke, D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 163196)
AUTHORS Gray, J.W., Collins, C., Kimmerly, W., Bondoc, M., Cheng, J.,
Connolly, K.S., Gunning, K.M., Davis, C.A., Kadner, K., Miguel, T.,
Pitluck, S., Pollard, M., Rojeski, H., Subramanian, S. and Martin, C.H.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source
1. 163196
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20"
/clone="BAC 99"
/note="LBNL H80"
4588..4709
/rpt_family="MIR"
6685..6813
repeat_region
/rpt_family="Tigger1"
complement(6838..7220)
repeat_region
/rpt_family="L1"
7175..7314
repeat_region
/rpt_family="L1"
7322..7441
repeat_region
/rpt_family="Tigger1"
complement(8695..9257)
repeat_region
/rpt_family="MLT2B2"
10121..10350
repeat_region
/rpt_family="Alu"
11057..11357
repeat_region


```
repeat_region /rpt_family="Alu"
13431. .13702
repeat_region /rpt_family="Alu"
15911. .15962
misc_feature /rpt_family="MIR"
16029. .16484
repeat_region /note="99% identity B54864 BAC end sequence"
16891. .17025
repeat_region /rpt_family="Alu"
17221. .17682
repeat_region /rpt_family="L1"
18500. .18756
misc_feature /rpt_family="L1"
18909. .19161
repeat_region /note="GRAIL 2 excellent exon, frame 1"
19640. .19943
repeat_region /rpt_family="Alu"
22025. .22314
misc_feature /rpt_family="Alu"
22903. .23020
repeat_region /note="GRAIL 2 excellent exon, frame 1"
23247. .23567
misc_feature /rpt_family="Alu"
24441. .24514
repeat_region /note="GRAIL 2 excellent exon, frame 0"
25506. .25525
repeat_region /note=" (T) 20"
/rpt_type=tandem
/rpt_unit="t"
repeat_region complement(25538. .25679)
repeat_region /rpt_family="MIR"
26920. .27577
repeat_region /rpt_family="LTR8"
28412. .28563
repeat_region /rpt_family="WER20"
28734. .29019
repeat_region /rpt_family="Alu"
30266. .30455
repeat_region /rpt_family="THE1"
30789. .30888
misc_feature /note="GRAIL 2 excellent exon, frame 2"
31841. .31972
misc_feature /note="GRAIL 2 excellent exon, frame 0"
32270. .32290
repeat_region /note=" (T) 21"
/rpt_type=tandem
/rpt_unit="t"
misc_feature 32702. .32860
misc_feature /note="GRAIL 2 excellent exon, frame 1"
33268. .33474
repeat_region /note="GRAIL 2 excellent exon, frame 0"
33836. .33863
repeat_region /note=" (GT)14"
/rpt_type=tandem
/rpt_unit="gt"
misc_feature 35158. .35265
repeat_region /note="GRAIL 2 excellent exon, frame 0"
41855. .42137
misc_feature /rpt_family="Alu"
43253. .43465
repeat_region /note="GRAIL 2 excellent exon, frame 1"
50487. .50600
repeat_region /rpt_family="L1"
50605. .50855
repeat_region /rpt_family="L1"
51461. .51822
repeat_region /rpt_family="THE1"
54343. .54856
repeat_region /rpt_family="LTR12"
55490. .55796
repeat_region /rpt_family="Alu"
55900. .56027
repeat_region /rpt_family="MIR"

complement(61007. .61096)
/note="GRAIL 2 excellent exon, frame 1"
join(61724. .61846,71254. .71319,75055. .75261)
/standard_name="af74a01"
/note="100% identity dbEST:AA644590"
complement(62247. .62746)
/rpt_family="L1"
repeat_region /rpt_family="Alu"
65258. .65541
repeat_region 65587. .65882
repeat_region /rpt_family="Alu"
68256. .68282
repeat_region /note=" (T) 27"
/rpt_type=tandem
/rpt_unit="t"
repeat_region 69077. .69495
repeat_region /rpt_family="Alu"
69537. .69618
repeat_region /rpt_family="Alu"
70542. .70721
repeat_region /rpt_family="MERS"
72072. .72255
repeat_region /rpt_family="MIR"
72260. .72340
repeat_region complement(72260. .72340)
74417. .74456
repeat_region /note=" (CT) 20"
/rpt_type=tandem
/rpt_unit="ct"
misc_feature 75057. .75320
prim_transcript /note="GRAIL 2 excellent exon, frame 2"
join(75199. .75323,84601. .84746)
/standard_name="zm96b07"
/note="96% identity dbEST:AA079436"
complement(77059. .77202)
repeat_region /rpt_family="MERS"
77269. .77551
repeat_region /rpt_family="Alu"
78685. .79043,88203. .88578)
prim_transcript join(78685. .79043,88203. .88578)
/standard_name="ym20b03"
/note="97% identity dbEST:H15040"
79526. .79590
prim_transcript /standard_name="nc17d10"
/note="81% identity dbEST:AA259187"
complement(81547. .81627)
repeat_region /rpt_family="MIR"
82700. .82865
repeat_region /rpt_family="Alu"
83191. .83347
repeat_region /rpt_family="L1"
83468. .83594)
repeat_region /rpt_family="MIR"
84144. .84436
repeat_region /rpt_family="Alu"
84402. .84774)
prim_transcript /standard_name="nl14b12"
/note="98% identity dbEST:AA554549"
complement(85065. .85162)
prim_transcript /standard_name="yw21h10"
/note="90% identity dbEST:H88501"
join(85165. .85589,86249. .86579,87354. .87476)
/standard_name="yg09d02"
/note="97% identity dbEST:R43020"
complement(86451. .86577)
misc_feature /note="GRAIL 2 excellent exon, frame 2"
87953. .88445)
prim_transcript /standard_name="zu90c12"
/note="99% identity dbEST:AA626177"
complement(89695. .89866)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
89959. .90146
repeat_region /rpt_family="MERS"
```



```
Query Match      33.2%; Score 1016.4; DB 9; Length 163196;
Best Local Similarity 99.1%; Pred. No. 2.7e-270;
Matches 1033; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1929 CCTGGGACCAACCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCCATCGGACC 1988
DB 83752 CCAGGACCAACCGGATGTTGGATGCTCAAGTGAACACAGACCCAGTATCCATCGGACC 83811
QY 1989 AGTTGGCAACCAAGTAACAAATCAGCACGGTTCACACAGGTTCTCTGCGCACCAAG 2048
DB 83812 AGTTGGCAATCCAAATGAACAAATCAGCACGGTTCACACAGGTTCTCTGCGCACCAAG 83871
QY 2049 ATGTGTTCTCTTACTTCCATCTCTCCCAACACGCTCCATGTATATATATCTCTGTATG 2108
DB 83872 ATGTGTTCTCTTACTTCCATCTCTCCCAACACGCTCCATGTATATATATCTCTGTATG 83931
QY 2109 GCCAGCAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTATT 2168
DB 83932 GCCAGCAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATATGAGGTGTATT 83991
QY 2169 ATTTAGTCTCTGGTCCAGTCTTTTCTGCGCAATTAACAGTAAGATGTTTAGCAGGTCA 2228
DB 83992 ATTTAGTCTCTGGTCCAGTCTTTTCTGCGCAATTAACAGTAAGATGTTTAGCAGGTCA 84051
QY 2229 CCTAGTGGGTCAGAGAGTGCATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGG 2288
DB 84052 CCTAGTGGGTCAGAGAGTGCATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGG 84111
QY 2289 TGTTCGGGTTAAGTGAATGAATGCGAGTGTGGCGGGCGGTGGTCTCGCCCTGTAA 2348
DB 84112 TGTTCGGGTTAAGTGAATGAATGCGAGTGTGGCGGGCGGTGGTCTCGCCCTGTAA 84171
QY 2349 TCTAGCACTTTGGGAGGCGGAGGAGGAGTGCATGATCCTGAGGTCAAGGATTTCAAGACTAG 2408
DB 84172 TCTAGCACTTTGGGAGGCGGAGGAGGAGTGCATGATCCTGAGGTCAAGGATTTCAAGACTAG 84231
QY 2409 CTTGGCCACATCATGAACCCGCTCTCTACTAATAATACAAAATTAAGGATTTAGCAGGATG 2468
DB 84232 CTTGGCCACATCATGAACCCGCTCTCTACTAATAATACAAAATTAAGGATTTAGCAGGATG 84291
QY 2469 GGCACACACCTGTAGTCTCCAGCTACTCGGGAGGCCAACCGACGAGAACCGTCTGTACCCA 2528
DB 84292 GGCACACACCTGTAGTCTCCAGCTACTCGGGAGGCCAACCGACGAGAACCGTCTGTACCCA 84351
QY 2529 GGAGGTGAGGTTGAGTGAAGCGGAGGTTGACCAATGACATTCACCTCCACCTGGGCGACAG 2588
DB 84352 GGAGGTGAGGTTGAGTGAAGCGGAGGTTGACCAATGACATTCACCTCCACCTGGGCGACAG 84411
QY 2589 CAAGATTCTATC--AAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2645
DB 84412 CAAGATTCTATCAAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 84471
QY 2646 CTAGAAGGAATTAAGCGTTGTAGTAAACCGGTGCTCATCTTAAGCTTGAAGAGGAGGAG 2705
DB 84472 CTAGAAGGAATTAAGCGTTGTAGTAAACCGGTGCTCATCTTAAGCTTGAAGAGGAGGAG 84531
QY 2706 ACAGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2765
DB 84532 ACAGAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 84591
QY 2766 GGTTCGCAATTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAGCGTTG 2825
DB 84592 GGTTCGCAATTCCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAGCGTTG 84651
QY 2826 TTTTTCAAATAGGTCCTTTGTAAGTTATTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2885
DB 84652 TTTTTCAAATAGGTCCTTTGTAAGTTATTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 84711
QY 2886 TTAATAATGATTTGGGAAACAAATTCCTTTTGAAGGCTCAGTGACAAACGCGAAGGATTACAA 2945
DB 84712 TTAATAATGATTTGGGAAACAAATTCCTTTTGAAGGCTCAGTGACAAACGCGAAGGATTACAA 84771
QY 2946 CTTAAAAAANAAAAAANAAAAA 2967
```

```
DB 84772 CTTAAAAAANAAAAAATATAATAA 84793

RESULT 7
LOCUS AY219233 2916 bp mRNA linear ROD 20-OCT-2003
DEFINITION Mus musculus novel amplified in breast cancer-1 mRNA, complete cds.
ACCESSION AY219233
VERSION AY219233.1 GI:29648617
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2916)
Beardley, D.I., Kowbel, D., Lataxes, T.A., Mannino, J.M., Xin, H.,
Kim, W.J., Collins, C. and Brown, K.D.
Characterization of the novel amplified in breast cancer-1 (NABCL1)
gene product
Exp. Cell Res. 290 (2), 402-413 (2003)
2230202
PUBMED 14567997
REFERENCE 2 (bases 1 to 2916)
AUTHORS Brown, K.
DIRECT SUBMISSION
SUBMITTED (14-JAN-2003) Biochem. and Molec. Biology, LSU Health
Sciences Center, 1901 Perdido Street, New Orleans, LA 70112, USA
LOCATION/Qualifiers
1. 2916
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
89..1990
/note="NABCL1; AIBC1; BCAS1"
/codon_start=1
/product="novel amplified in breast cancer-1"
/protein_id="AA088011.1"
/db_xref="GI:29648618"
/translation="MGQMVSPLRPGDQEHDPGADTKVTSNDECVQNGNVPVILSTR
IQHYEVDLGISSKDNVATSSPKTMEAOAVGDASGNLKGAKTKAPAAASHFTLTL
SRVPGRPGDQTDSSAERFDVSPSAAPENKDPSEHGALPVAAPQAAADPKTPGCP
EAKQOTLPATGLAPSPESAQAEPADQKDFGLNRFKLDKXGSEAPVNSQPEAKG
SEDEQATEAPAPVGNPHGVASGDIVDSORGQDVTLSYSVPGDPEVPTTKEDPQ
VVDITNSIMSFFTLVSPNKTETKDDPKDADKADSVCDGHAAGKMSQTSQAKS
KKKLDSPRLGLSPRLFRKRDSTNSPTTANLKSADKANFTPOETRGTKATKSGSP
PPSEPTSEGRDSEKAGPTLLPLGKLFWKSVKEDTLSTGABENAVCEBSEVTRL
EEVSSLTQVDLSEETOPEPTDVKVKEEKPKRTPLMAFLQMSVRSSEGIPIRSESN
VKDSSCOTNSVEKTPSPPEPAGTAKQNKETSSSKDKKSVDKKSATENSQKNGKQ
EVREPAFCVQPPPTVEANAMOTGDKTPKKSEKRRSGLGFLGLGPKRMSDAQVQDTPV
SIGPVGKSK"

ORIGIN
Query Match 29.8%; Score 912.8; DB 10; Length 2916;
Best Local Similarity 67.1%; Pred. No. 9.3e-242;
Matches 1564; Conservative 0; Mismatches 567; Indels 201; Gaps 11;

QY 163 CACGACTGGAGACCCAGGACCTGTGCAGGAGCAGCTCCGGGTGACACGAGGGGACTGA 222
DB 1 CAACCGCTGAAGACGAGGACCTTGAGCAGAACCACTGTGGAACTGTGGGGGACCCA 60
QY 223 AGATACTCCCAACAGGGCTCAGCAGGAGCAATGGGTAAACCAATAGTGTTCCTCCAAAGA 282
DB 61 AGATACTCCCAACCA--CTTACCAGGAACAATGGGAACCAAGATGAGCGTCCCACTAAGG 118
QY 283 GTTGAACACCAAGAGATGAACAGCAGCAGAGCTTACCAG-----GACACACGG 333
DB 119 CCTGGAGACCAAGGAGCAGCAGCAGGAGCAGACACTTGCACAGGTGACGTGACAGCAATGAG 178
QY 334 TCTGCTCTGAACAGGGGTTCCAGTGTGTGTGTCGACCCACACACAGTTCAGCAGCTTAGAGGAA 393
DB 179 TGTGTTTCAAGATGGGAACCCAGTTGTGCTGCTCCACTGTGTAAATTCACACTATGAGGAA 238
```

[illegible]

Db	1301	1361	1393	1420	1453	1479	1513	1537	1632	1597	1633	1692	1657	1749	1717	1794	1777	1854	1837	1914	1897	1974	1957	2034	2017	2094	2067	2154	2127	2214	2187	2274	2247	2326	2299
Db	AGCGCAAGAGAGGCGAGGCGCCACTTACTACCGCTGGAAAAGTTGTTTTGGAGAAG	TTTGGAGAAG	TCAGTTTAAAGAGAGACTCAGTCCCAAGAGTGGCGAGGAGAAATGTGTGTGTGAGTCAACCA	TCAGTTAAGGAGGATACATTTCCACAGGTGCAGAGGAGAAACGCGGTGTGTGAGTCAACCA	GTAGAGATTATAAGTCCAAAGGAAGTAGAATTCAGCTTTACAAACAGTGGACCTCAACGAA	GTAGAGACCGTAAAGCTTGGAGAAAGTAGAATTCAGCTTTACAACTGTGGATCTCAGTGA	GGAGATGCTGCACCTCAACCCACAGAGCGAACTCAAGAGAGAAAGAACCAACCAAGA	--AGAGACCCAGGCTTGAACCCACAGACGTAAGAAAGTCAAGAGAAAGAACCCCGGAAG	ACCTCTCTCATGTCGCTTTCTCAGACAAATGTCTAGTCAAAAGGGATGGAGGGAATCAACCCAC	ACCCCTCTCATGTCGCTTTCTCAGACAAATGTCTAGTGTGATCGAGCGAAGGGATCCCGCGC	TCAGAAAGAAATTAATGGGAAAGACTCAGCTGTCGCAACATCAGACTCCACAGAAAAAGACT	TCGAAAGAAAGTAACTGTAAGAGACTCAGCTGTCGCAACGTCACAACTCCGTGGAAAAAGACG	ATCACACCCGACAGCCTGAAACCAACAGGAGCACACAGAGGGTAAAGAGGGCTCC	CCCTCACCCCGACAGCCAGCGCGGGAACAGCTCAGAAAAACAGGAGACCTCTCTCC	-----TCGAAGGACAAGAAAGTCTCAGCAGCCGAGATGAAACAGCAGAGAGAGC	TCGAAGGACAAGAAATCAGTGGACAGAAAGTCTCAGCACTGAGAAACAGCAGCAGAGAGAC	AACAAGCAGGAAAGCCAAAGAACCCAGCCAGTGCACAGAGCAGCCACCGTGGACAGAAC	GGCAACAGAGAACTCAGAGAGCGCGCGTGTGTGCGAGCGCCACAGTGGAGGCAAC	TCATCGAGAAATGGGACAAAGCTCAGAAAGAGACCTGAGAGAGCGCAGCAGTCCCTTGGG	GCAATCAGACTGGGACAAAGACCCCAAGAAAGTCTGAGAGAGCGGACAGTCCCTCGGG	GGCTCTCTTAAAGGCTGGGACCAAGCGGATGTGTGATCTCAAGTGCACAAACAGACCCCA	GGCTTCTGAGGGCTGGACCAAGCGGATGTGATGCTCAGTGCAGTGCAGACGACCGC	GTATCCATCGGACAGTGGCAACCCAAAGTAAACAAATCAGACCGGTTCACCAACAGGTT	GTCTCCATCGGACAGTGGAAATCCAAAGTAAACAAATCACCAGTTCACCACTGGGTT	CTCTCGCACCAGATGTGTTCTCTTACTCCTATCTCTCCCAACACACGCTCCATGTAT	CTCTCTCCAC-----TCATCTACTGTCTCAGACCCACACTGTGTATATATTTTC	ATATTTCTCTGATGGCCAGCAATGAAATCTCTGCTAGAAATTAAGCCCGAGCTGTGTGA	TTCTTCTCTCTCTGATGAGCAATGAAATCTCTGCTTACAACTCACTGATGTAAAAAT	TATTGAGGTGATATTTTACGTCTCTGGTCCAGTCTTTCTGGCAATAACAGTAAAGAT	GATCGAGGTGATATTTAGTCTCTGGTCCAGTCTTCTCTGCAATTAAGTGTAAAAAT	GGTTTACGAGGTCACTAGTGGGTGAGAAAGATCGATGATCAACAGCAGGAAAGGAG	GGCTTAGCAGGTTGACCAAGTGTATCCAAAGGAGCTGGTGTGATGTCAGTGGGAGGAAG	GGAAATAGAGAAATGTTCTGGGTAAAGTGTGATGAAATGGCAGTGGTGGCGG	TAGGTAGATAAAGGCATTTGTGTAAAGTTATGAAAAAATGGCAGGTATCCAG	

RESULT 8

AC004501/c

LOCUS

DEFINITION

AC004501

Homo sapiens chromosome 20, BAC clone 121 (LBNL H144), complete sequence.

39507 bp

DNA

linear

PRI 20-JUN-1998

RESULT 8	AC004501/c	LOCUS	AC004501	39507 bp	DNA	linear	PRI 20-JUN-1998
DEFINITION		Homo sapiens chromosome 20, BAC clone 121 sequence.					(LBNL H144), complete

```

ACCESSION AC004501
VERSION AC004501.1 GI:2996638
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 39507)
Gray,J.W., Collins,C., Kimmerly,W., Bondoc,M., Cheng,J.,
Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T.,
Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
Sequencing of human chromosome 20.
TITLE
JOURNAL
REFERENCE
AUTHORS Unpublished
2 (bases 1 to 39507)
Ricke,D.O.
Large Scale Sequence Analysis and Annotation with the Sequence
Comparison Analysis (SCAN) System
JOURNAL
AUTHORS Unpublished
3 (bases 1 to 39507)
Gray,J.W., Collins,C., Kimmerly,W., Bondoc,M., Cheng,J.,
Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T.,
Pitluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
TITLE
JOURNAL Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
4 (bases 1 to 39507)
Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M.,
Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M.,
Rojeski,H., Subramanian,S. and Martin,C.H.
Direct Submission
TITLE
JOURNAL Submitted (30-MAR-1998) Human Genome Center, DOE Joint Genome
Institute, Lawrence Berkeley National Laboratory, MS 74-157,
Berkeley, CA 94720, U.S.A.
5 (bases 1 to 39507)
Ricke,D.O.
Direct Submission
TITLE
JOURNAL Submitted (20-JUN-1998) DOE Joint Genome Institute
Sequence submitted by:
COMMENT DOE Joint Genome Institute.
FEATURES
source
1. 39507
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20q13.2"
/clone="BAC121"
/note="LBNL H144"
misc_feature Complement (1022..1606)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region complement (1981..2192)
/rpt_family="MER20"
repeat_region 2230..2358
/rpt_family="MIR"
repeat_region 2889..2979
/rpt_family="MIR"
repeat_region 4477..7764
/rpt_family="L1"
repeat_region 7749..8061
/rpt_family="Alu"
repeat_region 8035..8054
/note="(A)20"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 8102..8880
/rpt_family="L1"
repeat_region 9213..9313
/rpt_family="MIR"
repeat_region complement (9732..10100)
/rpt_family="THE1b"
repeat_region complement (10253..10394)
/rpt_family="MIR"
repeat_region 13184..13974
/rpt_family="L1"
repeat_region 14234..14580
/rpt_family="Alu"
repeat_region 14544..14565
/note="(A)22"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 14875..15054
/rpt_family="L1"
repeat_region 15932..15964
/note="(TGA)11"
/rpt_type=tandem
/rpt_unit="tga"
repeat_region complement (16272..16455)
/rpt_family="MERS"
repeat_region 17190..17494
/rpt_family="Alu"
repeat_region 17547..17879
/rpt_family="Alu"
repeat_region 18872..19495
/rpt_family="Alu"
repeat_region 19163..19182
/note="(A)20"
/rpt_type=tandem
/rpt_unit="a"
repeat_region 19196..19484
/rpt_family="Alu"
misc_feature complement (20307..20596)
/note="83% identity AC003029"
misc_feature complement (20380..20726)
/note="84% identity B35601"
misc_feature complement (20639..20722)
/note="84% identity zx58h03.r1"
/db_xref="dbEST:AA203298"
repeat_region complement (22197..22354)
/rpt_family="Alu"
repeat_region complement (22435..22514)
/rpt_family="MIR"
repeat_region complement (22575..22996)
/rpt_family="MLT1"
repeat_region complement (24349..24689)
/rpt_family="Alu"
repeat_region 24361..24390
/note="(TTG)10"
/rpt_type=tandem
/rpt_unit="ttg"
repeat_region 24909..25022
/rpt_family="L1"
repeat_region 25103..25246
/rpt_family="MLT1a"
repeat_region 25335..25502
/rpt_family="MLT1"
repeat_region complement (26527..26896)
/rpt_family="THE1b"
repeat_region 27826..27889
/rpt_family="L1MD1"
repeat_region complement (27900..28255)
/rpt_family="THE1"
repeat_region 28248..29228
/rpt_family="L1"
repeat_region complement (29552..29682)
/rpt_family="MIR"
repeat_region complement (29955..30101)
/rpt_family="MIR"
repeat_region 30224..30576
/rpt_family="MLT1"
misc_feature complement (30651..30788)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 31228..31247
/note="(T)20"
/rpt_type=tandem
/rpt_unit="t"

```


*	1	1374: contig of 1374 bp in length
*	1375	1474: gap of 100 bp
*	1475	3271: contig of 1797 bp in length
*	3272	3371: gap of 100 bp
*	3372	5701: contig of 2330 bp in length
*	5702	5801: gap of 100 bp
*	5802	8678: contig of 2877 bp in length
*	8679	8778: gap of 100 bp
*	8779	11405: contig of 2627 bp in length
*	11406	11505: gap of 100 bp
*	11506	15628: contig of 4123 bp in length
*	15629	15728: gap of 100 bp
*	15729	18099: contig of 2371 bp in length
*	18100	18199: gap of 100 bp
*	18200	21769: contig of 3570 bp in length
*	21770	21869: gap of 100 bp
*	21870	25732: contig of 3863 bp in length
*	25733	25832: gap of 100 bp
*	25833	31618: contig of 5786 bp in length
*	31619	31718: gap of 100 bp
*	31719	36866: contig of 5148 bp in length
*	36867	36966: gap of 100 bp
*	36967	46305: contig of 9339 bp in length
*	4306	46405: gap of 100 bp
*	46406	58028: contig of 11623 bp in length
*	58029	58128: gap of 100 bp
*	58129	72476: contig of 14348 bp in length
*	72477	72576: gap of 100 bp
*	72577	86507: contig of 16031 bp in length
*	86608	88707: gap of 100 bp
*	88708	102833: contig of 14126 bp in length
*	102834	102933: gap of 100 bp
*	102934	127147: contig of 24214 bp in length
*	127148	127247: gap of 100 bp
*	127248	150561: contig of 23314 bp in length
*	150562	150661: gap of 100 bp
URES	150662	181313: contig of 30652 bp in length.
source		Location/Qualifiers
		1..181313
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="20"
		/map="20"
		/clone="RP11-469D15"
		/clone.lib="RPC1-11 Human Male BAC"
misc_feature		1..1374
		/note="assembly_fragment"
misc_feature		1475..3271
		/note="assembly_fragment"
misc_feature		3372..5701
		/note="assembly_fragment"
misc_feature		5802..8678
		/note="assembly_fragment"
misc_feature		8779..11405
		/note="assembly_fragment"
misc_feature		11506..15628
		/note="assembly_fragment"
misc_feature		15729..18099
		/note="assembly_fragment
		clone_end:T7
		vector_side:left"
misc_feature		18200..21769
		/note="assembly_fragment"
misc_feature		21870..25732
		/note="assembly_fragment"
misc_feature		25833..31618
		/note="assembly_fragment"
misc_feature		31719..36866
		/note="assembly_fragment"
misc_feature		36967..46305
		/note="assembly_fragment"
misc_feature		46406..58028

[illegible][illegible]


```
RESULT 11
AX150111
LOCUS AX150111
DEFINITION Sequence 86 from Patent WO0136685.
ACCESSION AX150111
VERSION AX150111.1 GI:14348139
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Kroes, R.A., Moskal, J.R. and Yamamoto, H.
TITLE Differential gene expression in cancer
JOURNAL Patent: WO 0136685-A 86 25-MAY-2001;
NYXIS Neurotherapies, Inc. (US)
FEATURES
Location/Qualifiers
source
1..488
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 15.3%; Score 470.4; DB 6; Length 488;
Best Local Similarity 99.8%; Pred. No. 5.6e-119;
Matches 471; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2142 CCGAGCTGTGTATTCAGGCTGTATTTACGCTCTCGTCCAGTCTTTCTGGCAA 2201
Db 15 CCGAGCTGTGTATTCAGGCTGTATTTACGCTCTCGTCCAGTCTTTCTGGCAA 74
QY 2202 TAACAGTAAAGATGTTTACGAGTCACTAGTTGGTCCAGAGTGCATGCCAA 2261
Db 75 TAACAGTAAAGATGTTTACGAGTCACTAGTTGGTCCAGAGTGCATGCCAA 134
QY 2262 CAGAGAAAGGGAGGAATAGAGGATGTTTCGGTTAAGTCATCAAAATGCCAGTGTG 2321
Db 135 CAGAGAAAGGGAGGAATAGAGGATGTTTCGGTTAAGTCATCAAAATGCCAGTGTG 194
QY 2322 GCCGGGCGTGTGGCTCTCGCTCTAGTCTCAGACACTTTGGAGCCGAGGAGTGGAT 2381
Db 195 GCCGGGCGTGTGGCTCTCGCTCTAGTCTCAGACACTTTGGAGCCGAGGAGTGGAT 254
QY 2382 CACTTGAGTTCAGAGTTCAAGACTAGCTGGCCAACTATATGAACCCCGTCTCTACTA 2441
Db 255 CACTTGAGTTCAGAGTTCAAGACTAGCTGGCCAACTATATGAACCCCGTCTCTACTA 314
QY 2442 AAAATACAAAATTAGCCAGGATGTTGGCAGACACACTGTAGTCCAGCTACTCGGGAGC 2501
Db 315 AAAATACAAAATTAGCCAGGATGTTGGCAGACACACTGTAGTCCAGCTACTCGGGAGC 374
QY 2502 CCAAGCCAGGAGAACCGCTTGTACCCAGAGGTGGAGGTTCAGTGGCCGAAAGTTGCAC 2561
Db 375 CCAAGCCAGGAGAACCGCTTGTACCCAGAGGTGGAGGTTCAGTGGCCGAAAGTTGCAC 434
QY 2562 CATTGCATCCACCTCGGGCCAGAGCAAGATTCTATCAAAAAAAGG 2613
Db 435 CATTGCATCCACCTCGGGCCAGAGCAAGATTCTATCAAAAAAAGG 486
RESULT 12
AL935134/c
LOCUS AL935134
DEFINITION Mouse DNA sequence from clone RP23-321D1 on chromosome 2, complete
sequence.
ACCESSION AL935134
VERSION AL935134.10 GI:32400093
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

1 (bases 1 to 105616)

Humphries, M.

Direct Submission

Submitted (01-JUL-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Jul 1, 2003 this sequence version replaced gi:2988596.

Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun map have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em, EMBL; Sw, SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information
on the WormPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-321D1 is

from the RP23-23 Mouse BAC Library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES
source

1..105616
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-321D1"
/clone_lib="RPCI-23"

ORIGIN

Query Match 9.9%; Score 303.6; DB 10; Length 105616;
Best Local Similarity 72.1%; Pred. No. 2.4e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

QY 393 AGTCGACTTGGGATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGCAAC 452

Db 9080 AGTCGACTTGGGATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCTTCCCGGAGCAAC 9021

QY 453 GGAGATAAGTCTGTTCGGATGCCAAGGAAAGAACTCTGGGAAAGAGGCCAACCCGA 512

Db 9020 GGAGGCCAGCTGTGGCGATGCCGGAAGAAAGAACTCTGGGAAAGAGGCCAACCCGA 8961

QY 513 GGCAACGAGTCTGTAATCTCGTTTTTTTCTTGTATGCTCTCTCGGCTGTACCGAGCGTAC 572

Db 8960 GGCAACGAGTCTGTAATCTCGTTTTTTTCTTGTATGCTCTCTCGGCTGTACCGAGCGTAC 8901

QY 573 CGGAGACGAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTCAGTCAATAA 632

Db 8900 CGGAGACGAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTCAGTCAATAA 8841

QY 633 AGCTCCAGGAAACCAAGACCAAGTGGAGCTGGCAGCTTCCGCTGGCAGTGGACCGG 692

Db 8840 AGCTCCAGGAAACCAAGACCAAGTGGAGCTGGCAGCTTCCGCTGGCAGTGGACCGG 8781


```

misc_feature      13023..14162
                  /note="wgs_contig"
misc_feature      15743..16838
                  /note="wgs_contig"

ORIGIN
Query Match      9.9%; Score 303.6; DB 2; Length 168876;
Best Local Similarity 72.1%; Pred. No. 2.6e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTCCCGGAGACAC 452
Db 58999 AGTCGATTGGGAATCAACAGCTTGAAGGATAATGTGCCGCTTCTTCCCGGAGACAA 59058

Qy 453 GGAGATAAGTGTCTGTTCGGATGCCAACGGAAGAATCTTGGGAAGAGGCCAACCCGA 512
Db 59059 GGAAGCTCAGCTGTGGCTGATGCGACGGAAGAATCTTGGGAAGGATCCNAGCCCA 59118

Qy 513 GGACACAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCTGTACAGAGACGTAC 572
Db 59119 AGCACACAGCTGTAGTCCCAATTTTTTCTTGACACTCTCTCGGCTGTACAGAGACGTCC 59178

Qy 573 CGGAGACCAAGCCGACAGATTCCCTTGGATCAGTGAAGCTTGATGTCAGTCCCATAA 632
Db 59179 CGGGGACCAAGGCACAGATTCTATCGGCTGCATCGGGAGGCTTGATGTCAGCCCAAGGCAC 59238

Qy 633 AGCTCCAGCGCAACAAAGACCCAAAGTCAGAGCTGGACACTTCCGCTGGCAGCTGGACCGGG 692
Db 59239 GGGCTCTGAGACAAAGACCCGNGTGACACGGGGCACTTCCGCTGGCAGCTGCACACAG 59298

Qy 693 GCAGGACACAGATAAAACCCCAAGGGCACGCCCGCGGCCCAAGACAAGTCTCTCTGCGCG 752
Db 59299 GCGCGCTCCAGATAAAACCCCAAGGGTGCACTGAGGCCAAGCAGCAGACCTCTCTGCGCAC 59358

Qy 753 CAGGGATCCACAGCTTCTCCACCTGAGACAGGGGAGCAGAGGAGAGAGCTCCCTCCAA 812
Db 59359 CAGCCACTGGCACTTCCACCTGAGTCCGAGC-----AGAAAGCCCGTCCAG 59409

Qy 813 GCCCAAGGACTCCAGAGCTTTTTTGACAAAATCTTTCAAGCTGGACAAGGACAGGAAAAGGT 872
Db 59410 GCCCAGGATTTGAGCTTTTGAACAGACTCTTTAACTGGACACAGGAGAGAAAGCGC 59469

Qy 873 GCCAGGTGACAGCCAAAC---AGGAAGCCAAAGGGCAGAGCATCAAGACAAGGTGGATGA 929
Db 59470 GCGCGCTGACAGCGCCCTGAAAGAGGGAAAGGTCCTCAAGGCCAAGAACCGGCCATAGA 59529

Qy 930 GGTTCCTGGCTTATCAGGCGAGTCCGATGATGCTCCCTGAGGGAAGGACATAGT 983
Db 59530 GGCTCCTGCTGTACAGGGAATCTCTATGGTGTCTCTGAGGGAAGGTAAAGTGT 59583

RESULT 14
AC094145
LOCUS      232333 bp DNA linear HTG 09-MAY-2003
DEFINITION Rattus norvegicus clone CH230-2L13, WORKING DRAFT SEQUENCE, 6
            unordered pieces.
ACCESSION  AC094145
VERSION    6 GI:30467887
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE
1 (bases 1 to 232333)
Muzny,D.,Marle., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Blawalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,J., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,

```

```
----- Project Information
Center project name: GABR
Center clone name: CH230-2L13
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 207295 bases at least Q40
Consensus quality: 210461 bases at least Q30
Consensus quality: 212871 bases at least Q20
Estimated insert size: 217628; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 6 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
* 1 89147: contig of 89147 bp in length
* 89148 89247: gap of unknown length
* 89248 99513: contig of 10266 bp in length
* 99514 99613: gap of unknown length
* 99614 113910: contig of 14297 bp in length
* 113911 114010: gap of unknown length
* 114011 230487: contig of 116477 bp in length
* 230488 230587: gap of unknown length
* 230588 232001: contig of 1414 bp in length
* 232002 232101: gap of unknown length
* 232102 233233: contig of 1132 bp in length.
* Location/Qualifiers
  1. 233233
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-2L13"
  misc_feature
    1. 1412
      /note="wgs_contig"

FEATURES
             source
1. 233233
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10116"
   /clone="CH230-2L13"
   misc_feature
     1. 1412
       /note="wgs_contig"

ORIGIN
Query Match          9.9%; Score 303.6; DB 2; Length 233233;
Best Local Similarity 72.1%; Pred. No. 2.7e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy 393 AGTGACATTGGGATAAGTGTCAAGACGGATAATGTGCGCACTTCTTCCCGGAGACAAAC 452
Db 215357 AGTCGATTGGGAATCAACACAGCTTGAAGGATAATGTGCGCGCTTCTTCCCGGAGACAAT 215416
Qy 453 GGAGATAAGTGTGTTCGGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCAAACCCGA 512
Db 215417 GGAGCTCACGCTGTGGCTGTATGCCAGCGGAAGAAAGAAATCTTGGGAAAGATCAAGCCCA 215476
Qy 513 GGCACACAGCTGCTAAATCTGTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGCPAC 572
Db 215477 AGCACACAGCTGCTAGTCCCATTTTCTTGACACTCTCTCGGCTGTACAGGACGCTCC 215536
Qy 573 CGGAGACCAAGCCGAGATTATCCCTTGGATCAGTGAAGCTTGATGTGACTCAATAA 632
Db 215537 CGGGACCAAGGCACAGATTATCCGCTGTCATCGGGGAGGTTGATGTACGCCACGGCAC 215596
Qy 633 AGCTCCAGCAACAAAGACCAAGTGAGAGCTGGACACTTCGGTGGCAGCTGGACCGGG 692
Db 215597 GCGGCTTGAGAAACAAAGACCCGAGTGAGCACGGGCACTTTCGGTGGCAGCTGCACACAGG 215656
Qy 693 GCAGGACACAGATAAAACCCAGGCGACGCCGCCCAAGCAAAAGGTCTCTCTGCGCGC 752
Db 215657 GCGGCTCCAGATAAAACCCAGGCGTGCATCGAGGCCAAGCAGCAGACCTTCCCTGCCAC 215716
Qy 753 CAGGATCCAGCTTCTCCACCTGACACAGGGGAGCAGGAGAGAGCTCCCTCCAA 812
Db 215717 CAGCCCACTGGCACCTTCAACCACTTGAGTCCAGGC-----AGAAAGCCCGTCCAG 215767

----- Project Information
Center project name: GABR
Center clone name: CH230-2L13
----- Summary Statistics
Assembly program: Atlas;
Consensus quality: 207295 bases at least Q40
Consensus quality: 210461 bases at least Q30
Consensus quality: 212871 bases at least Q20
Estimated insert size: 217628; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 6 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
* 1 89147: contig of 89147 bp in length
* 89148 89247: gap of unknown length
* 89248 99513: contig of 10266 bp in length
* 99514 99613: gap of unknown length
* 99614 113910: contig of 14297 bp in length
* 113911 114010: gap of unknown length
* 114011 230487: contig of 116477 bp in length
* 230488 230587: gap of unknown length
* 230588 232001: contig of 1414 bp in length
* 232002 232101: gap of unknown length
* 232102 233233: contig of 1132 bp in length.
* Location/Qualifiers
  1. 233233
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-2L13"
  misc_feature
    1. 1412
      /note="wgs_contig"

FEATURES
             source
1. 233233
   /organism="Rattus norvegicus"
   /mol_type="genomic DNA"
   /db_xref="taxon:10116"
   /clone="CH230-2L13"
   misc_feature
     1. 1412
       /note="wgs_contig"

ORIGIN
Query Match          9.9%; Score 303.6; DB 2; Length 233233;
Best Local Similarity 72.1%; Pred. No. 2.7e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy 393 AGTGACATTGGGATAAGTGTCAAGACGGATAATGTGCGCACTTCTTCCCGGAGACAAAC 452
Db 215357 AGTCGATTGGGAATCAACACAGCTTGAAGGATAATGTGCGCGCTTCTTCCCGGAGACAAT 215416
Qy 453 GGAGATAAGTGTGTTCGGGATGCCAAGGAAAGAAATCTTGGGAAAGAGGCCAAACCCGA 512
Db 215417 GGAGCTCACGCTGTGGCTGTATGCCAGCGGAAGAAAGAAATCTTGGGAAAGATCAAGCCCA 215476
Qy 513 GGCACACAGCTGCTAAATCTGTTTTTCTTGATGCTCTCTCGGCTGTACAGGACGCPAC 572
Db 215477 AGCACACAGCTGCTAGTCCCATTTTCTTGACACTCTCTCGGCTGTACAGGACGCTCC 215536
Qy 573 CGGAGACCAAGCCGAGATTATCCCTTGGATCAGTGAAGCTTGATGTGACTCAATAA 632
Db 215537 CGGGACCAAGGCACAGATTATCCGCTGTCATCGGGGAGGTTGATGTACGCCACGGCAC 215596
Qy 633 AGCTCCAGCAACAAAGACCAAGTGAGAGCTGGACACTTCGGTGGCAGCTGGACCGGG 692
Db 215597 GCGGCTTGAGAAACAAAGACCCGAGTGAGCACGGGCACTTTCGGTGGCAGCTGCACACAGG 215656
Qy 693 GCAGGACACAGATAAAACCCAGGCGACGCCGCCCAAGCAAAAGGTCTCTCTGCGCGC 752
Db 215657 GCGGCTCCAGATAAAACCCAGGCGTGCATCGAGGCCAAGCAGCAGACCTTCCCTGCCAC 215716
Qy 753 CAGGATCCAGCTTCTCCACCTGACACAGGGGAGCAGGAGAGAGCTCCCTCCAA 812
Db 215717 CAGCCCACTGGCACCTTCAACCACTTGAGTCCAGGC-----AGAAAGCCCGTCCAG 215767

RESULT 15
AC084066
LOCUS       AC084066                235411 bp    DNA    linear    HTG 12-OCT-2000
DEFINITION Mus musculus clone RP23-321D1, *** SEQUENCING IN PROGRESS ***, 29
            unordered pieces.
ACCESSION   AC084066
VERSION     AC084066.1 GI:10799415
KEYWORDS    HTG; HTGS PHASE1.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 235411)
AUTHORS     DOE Joint Genome Institute.
TITLE       Sequencing of Mouse
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 235411)
AUTHORS     DOE Joint Genome Institute.
TITLE       Direct Submision
JOURNAL     Submitted (12-OCT-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     -----Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -----
Project Information
Center Project Name: 2351294
Center clone name: RPCI-23_321D1
-----
Summary Statistics
Consensus quality: 214207 bases at least Q40
Consensus quality: 223053 bases at least Q30
Consensus quality: 225208 bases at least Q20
Estimated insert size: 200000; pulse field gel estimation
Estimated insert size: 232611; sum-of-contigs estimation
Quality coverage: 9.94 in Q20 bases; pulse field gel estimation
Quality coverage: 11.56 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
  consists of 29 contigs. The true order of the pieces
  is not known and their order in this sequence record is
  arbitrary. Gaps between the contigs are represented as
  runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
  as soon as it is available and the accession number will
  be preserved.
* 1 1210: contig of 1210 bp in length
* 1310: gap of unknown length
* 1311 1310: contig of 1930 bp in length
* 3240: contig of 1930 bp in length
* 3241 3340: gap of unknown length
* 3341 4504: contig of 1164 bp in length
* 4504 4604: gap of unknown length
* 4604 4605: gap of unknown length
* 4605 6172: contig of 1568 bp in length
* 6173 6272: gap of unknown length
* 6272 7354: contig of 1082 bp in length
* 7354 7454: gap of unknown length
* 7454 8625: contig of 1171 bp in length
* 8625 8726: gap of unknown length
* 8726 10114: contig of 1389 bp in length
```

```
* 10115 10214: gap of unknown length
* 10215 12091: contig of 1877 bp in length
* 12092 12191: gap of unknown length
* 12192 14113: contig of 1922 bp in length
* 14114 14213: gap of unknown length
* 14214 15286: contig of 1073 bp in length
* 15287 15386: gap of unknown length
* 15387 17839: contig of 2453 bp in length
* 17840 17939: gap of unknown length
* 17940 20266: contig of 2327 bp in length
* 20267 20366: gap of unknown length
* 20367 22001: contig of 1635 bp in length
* 22002 22101: gap of unknown length
* 22102 24307: contig of 2206 bp in length
* 24308 24407: gap of unknown length
* 24408 26405: contig of 1998 bp in length
* 26406 26505: gap of unknown length
* 26506 29691: contig of 3186 bp in length
* 29692 29791: gap of unknown length
* 29792 34705: contig of 4914 bp in length
* 34706 34805: gap of unknown length
* 34806 39749: contig of 4944 bp in length
* 39750 39849: gap of unknown length
* 39850 45296: contig of 5447 bp in length
* 45297 45396: gap of unknown length
* 45397 51476: contig of 6080 bp in length
* 51477 51576: gap of unknown length
* 51577 59008: contig of 7432 bp in length
* 59009 59108: gap of unknown length
* 59109 66218: contig of 7110 bp in length
* 66219 66318: gap of unknown length
* 66319 76778: contig of 10460 bp in length
* 76779 76878: gap of unknown length
* 76879 80098: contig of 21220 bp in length
* 80099 98198: gap of unknown length
* 98199 113987: contig of 15789 bp in length
* 113988 114087: gap of unknown length
* 114088 132115: contig of 18028 bp in length
* 132116 132215: gap of unknown length
* 132216 160184: contig of 27969 bp in length
* 160185 160284: gap of unknown length
* 160285 186698: contig of 26414 bp in length
* 186699 186798: gap of unknown length
* 186799 235411: contig of 48613 bp in length.
FEATURES
    source
    1..235411
    /organism="Mus musculus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10090"
    /clone="RP23-321D1"
    /clone_lib="RPC1 mouse BAC library 23"
```

ORIGIN

```
Query Match          9.9%; Score 303.6; DB 2; Length 235411;
Best Local Similarity 72.1%; Pred. No. 2.7e-72;
Matches 428; Conservative 0; Mismatches 154; Indels 12; Gaps 2;

Qy 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTCCCGGAGACAAC 452
Db 84020 AGTCGACTTGGGAATCAGCAGCTCCAGAGATATGTGGCTACTTCTCCCGGAGACAAT 84079

Qy 453 GGAGATAAGTGTGTGGGATGCCAACGGAAAGAAATCTGGGAAGAGGCCAACCCGA 512
Db 84080 GGAGGCCCGAGGTGTGGGCGATGCCAGCGGAAGAATCTGGGAAGAGGCCAACCA 84139

Qy 513 GGCACCACTGCTAAATCTCGTTTCTTCTGATGCTCTCTGGCGCTGTACGAGGACGTAC 572
Db 84140 GGCACCACTGCTAGATCCCAATTTTCTTGACACTCTCTCGGCTGTACGAGGACGTCC 84199

Qy 573 CGGAGACCAAGCCGAGATTCATCCCTTGGGATCAGTGAAGCTTGATGTGAGTCCCAATAA 632
Db 84200 CGGGACCAAGGCACGGATTCTATCGGCTGCTATCAGGAGGTTGTATGTACGCCCAAGCGC 84259
```

```
Qy 633 AGTCCAGCGAACAAAGACCCCAAGTGAAGCTGGACACTTCCGGTGGCAGCTGGACCGG 692
Db 84260 AGCGCTGAGAAACAAAGACCCCAAGTGAAGCAGCGGGCACTTCCGGTGGCAGCTGGACCCAGG 84319

Qy 693 GCAGGACACAGATAAAACCCCGAGGGCAGCGCCCGGCCCAAGACAAGGTCTCTCTGCGCG 752
Db 84320 GCAGGCTCCAGATAAACCCTCCAGGGTGCCTCCGAGGCCAAGCAGCAGACCTCTCCCTGCCAC 84379

Qy 753 CAGGGATCCCAAGCTTCTCCCACTGAGACAGGGGGAGCAGGAGAGAAAGTCTCCCTCCAA 812
Db 84380 CGGCCCATTTGCAACCTCTCACCACTGAGTCCAGGC-----AGAAGCCCGGCCCA 84430

Qy 813 GCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAGGGACAGGAAAGGT 872
Db 84431 GGACAAGGATTTCCGGCTTTTGTAAACAGATTCTTTAACTGGACAAAGGAAGAAAGCGC 84490

Qy 873 GCCAGGTGACACAGCCAAC---AGGAAGCCCAAGAGGGCAGAGCATCAAGACAAGGTGGATGA 929
Db 84491 GCCGTGAACAGCAGCCCAAGAAAGCGAAAGGCTCGAAGACCCAGAACAGGCCACAGA 84550

Qy 930 GGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCTCTGACGGGAGGACATAGT 983
Db 84551 GGCTCCTGCCGTGCCAGGGAATCCCATGGTGTCTCTGACGGGAGGTAAGTGT 84604
```

Search completed: July 4, 2005, 01:47:00
Job time : 13007 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 1566.51 Seconds

(without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGACACGCTATGACCATGA.....GGTACCAATTGCGCCTATA 3066

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3066	100.0	3066	2	AAV09025 Homo sapi
2	2828.8	92.3	2418	12	ADQ23347 Human sof
3	2813	91.7	2813	5	AA81115 DNA encod
4	2813	91.7	2813	6	ADQ38092 Human BCA
5	2813	91.7	2813	12	ADJ75127 Marker ge
6	2813	91.7	2813	12	ADQ18993 Human sof
7	2813	91.7	2813	13	ADQ18993 Breast ca
8	1929.4	62.9	2020	8	ACC90605 Human CGD
9	1567	51.1	2105	4	AA812521 Gene #14
10	1567	51.1	2105	10	ABQ77258 Human ten
11	1016.4	33.2	9889	5	ABA17779 Human ner
12	584	19.0	1504	8	AB236099 Human sec
c 13	582.6	19.0	644	12	ACH89479 Human gen
c 14	470.4	15.3	488	4	AAH50757 Human tum
c 15	332.8	10.9	576	12	ACH70719 Human gen
c 16	328	10.7	372	10	ADF85762 Human ade
c 17	313	10.2	382	4	AA180677 Human pol
c 18	297	9.7	506	12	ACH75758 Human gen
c 19	289.2	9.4	540	13	ADQ53268 Novel can
20	265	8.6	265	4	AA152231 Probe #20

21	265	8.6	265	4	AAK46334 Human bon
22	265	8.6	265	4	AAK20263 Human bra
23	265	8.6	265	4	ABS46058 Human liv
24	265	8.6	265	6	ABS20652 Human gen
25	260.4	8.5	495	4	AAI39177 Probe #78
26	260.4	8.5	495	4	AAK33393 Human bon
27	260.4	8.5	495	4	AAK07595 Human bra
28	260.4	8.5	495	4	ABS33149 Human liv
29	260.4	8.5	495	6	ABS08233 Human gen
30	247	8.1	349	2	AAV88411 EST clone
c 31	242.6	7.9	6096	6	ABK92513 Human pro
c 32	240.8	7.9	214520	10	ADL13471 Osteoartrh
33	238	7.8	297	3	AAK10324 Human sec
c 34	236	7.7	181684	11	ACN44374 Human gen
35	233.8	7.6	68355	8	ACF62737 Cancer ba
36	233.8	7.6	68355	8	ADB20852 MRP1 base
37	233.8	7.6	68355	10	ADB87941 Human UGT
38	233.8	7.6	68355	10	ADB96924 Human MDR
39	233.8	7.6	68355	10	ADB92115 Human MDR
40	233.8	7.6	186591	8	ACF62750 Cancer ba
41	233.8	7.6	186591	8	ADB20869 MRP1 base
42	233.8	7.6	186591	10	ADB87958 Human UGT
43	233.8	7.6	186591	10	ADB96941 Human MDR
44	233.8	7.6	186591	10	ADB92132 Human MDR
c 45	233.8	7.6	208648	8	ACF62735 Cancer ba

ALIGNMENTS

RESULT 1
AAV09025 AAV09025 standard; cDNA; 3066 BP.
XX
XX
XX AAV09025;
XX
XX 21-JUL-1998 (first entry)
XX
XX Homo sapiens 20ql3 amplicon lbl transcript.
XX
XX 20ql3 amplicon; chromosome 20; tumour; detection;
XX
XX chromosomal abnormalities; probe; gene therapy; antisense inhibition;
XX
XX treatment; age-related macular degeneration; retinitis pigmentataion;
XX
XX Leber's congenital amaurosis; ds.
XX
XX Homo sapiens.
XX
XX WO9802539-A1.
XX
XX PD 22-JAN-1998.
XX
XX PF 15-JUL-1997; 97WO-US012343.
XX
XX PR 15-JUL-1996; 96US-00680395.
XX
XX PR 16-OCT-1996; 96US-00731499.
XX
XX PR 17-JAN-1997; 97US-00785532.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX
XX WPI; 1998-110587/10.
XX
XX New sequences from the 20ql3 amplicon - used for detecting chromosomal
XX
XX abnormalities, particularly tumours, and for developing products for
XX
XX treating diseases.
XX
XX Claim 1; Page 69-70; 91pp; English.
XX
XX The sequence is that of a cDNA sequence lbl, which was isolated from the
XX
XX 20ql3 amplicon, it is overexpressed in numerous breast cancer cell lines
XX
XX and some primary tumours. It can be used as a probe for the detection of
XX
XX chromosomal abnormalities at 20ql3. It and other sequences isolated from

CC the 20q13 amplicon are consistently amplified in primary tumours. These
CC sequences are useful as probes or as probe targets for monitoring the
CC relative copy number of corresponding sequences from a biological sample
CC such as tumour cells. The sequences can also be used in therapeutic
CC applications for modulating the expression of the endogenous gene or the
CC activity of the gene product. Examples of therapeutic approaches include
CC antisense inhibition of gene expression, gene therapy, and monoclonal
CC antibodies that specifically bind the gene products. The products can
CC also be used in the treatment of other diseases, e.g. age-related macular
CC degeneration, Leber's congenital amaurosis and retinitis pigmentation
XX
SQ Sequence 3066 BP; 944 A; 741 C; 818 G; 563 T; 0 U; 0 Other;

Query Match 100.0%; Score 3066; DB 2; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGAACAGCTATGACCATGATTAAGCCAAAGCTCGAAATTAACCTCTCAATAAGGGAACAA 60
Db 1 GGAACAGCTATGACCATGATTAAGCCAAAGCTCGAAATTAACCTCTCAATAAGGGAACAA 60
Qy 61 AAGCTGGAGCTCCACCGGCTGGCGGCGCTCTAGAACTAGTGGATCCCGGGCTGCAG 120
Db 61 AAGCTGGAGCTCCACCGGCTGGCGGCGCTCTAGAACTAGTGGATCCCGGGCTGCAG 120
Qy 121 GAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGCGACGCACTGGAGACCCAG 180
Db 121 GAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGCGACGCACTGGAGACCCAG 180
Qy 181 GACCTGTGAGGAGAGCTCCCGGTGACAGAGGGAGCTGAAGATCTCCACAGGGGC 240
Db 181 GACCTGTGAGGAGAGCTCCCGGTGACAGAGGGAGCTGAAGATCTCCACAGGGGC 240
Qy 241 TCACAGAGCAATGGGTAAACCAATGAGTCTCCCAAGAGTTGAAGACCAAGAGAT 300
Db 241 TCACAGAGCAATGGGTAAACCAATGAGTCTCCCAAGAGTTGAAGACCAAGAGAT 300
Qy 301 GAAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Db 301 GAAACAGAGCAGAGACTTACAGGACAAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Qy 361 GTGTGACCCCAACAGTTCAGCACTTAGAGAAAGTGCACCTTGGGAATTAAGTGTCAAGACG 420
Db 361 GTGTGACCCCAACAGTTCAGCACTTAGAGAAAGTGCACCTTGGGAATTAAGTGTCAAGACG 420
Qy 421 GATATGTGGCCACTTCTTCCCGGACACAGGATAGTGTCTTGGGATGCCAAC 480
Db 421 GATATGTGGCCACTTCTTCCCGGACACAGGATAGTGTCTTGGGATGCCAAC 480
Qy 481 GGAAGAAATCTTGGGAAGAGGCCAAACCCGAGGCACAGCTGTCTAAATCTCGTTTTTTC 540
Db 481 GGAAGAAATCTTGGGAAGAGGCCAAACCCGAGGCACAGCTGTCTAAATCTCGTTTTTTC 540
Qy 541 TTGATGTCTCTCGGCTGTACAGGACGTAACCGGAGACCAAGCCGCGAGATTCATCCCTT 600
Db 541 TTGATGTCTCTCGGCTGTACAGGACGTAACCGGAGACCAAGCCGCGAGATTCATCCCTT 600
Qy 601 GGATCAGTGAAGCTTGATGTGAGTCCAAATAGCTCCAGGACCAAGACCCCAAGTGAG 660
Db 601 GGATCAGTGAAGCTTGATGTGAGTCCAAATAGCTCCAGGACCAAGACCCCAAGTGAG 660
Qy 661 AGCTGGACACTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAATCCAGGGCAC 720
Db 661 AGCTGGACACTTCCGGTGGCAGCTGGACCGGGCAGGACACAGATAAATCCAGGGCAC 720
Qy 721 GCCCGGCCCAAGACAGGCTCTCTGCGGCGAGGGATCCACGCTTCTCCACCTGAG 780
Db 721 GCCCGGCCCAAGACAGGCTCTCTGCGGCGAGGGATCCACGCTTCTCCACCTGAG 780
Qy 781 ACAGGGGAGCAGGAGAGAGCTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAA 840
Db 781 ACAGGGGAGCAGGAGAGAGAGCTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAA 840

Qy 841 TTCTTCAAGCTGGACAAAGGACAGGAAAAGGTGCCAGGTGACAGCCCAACAGGAAGCCAAG 900
Db 841 TTCTTCAAGCTGGACAAAGGACAGGAAAAGGTGCCAGGTGACAGCCCAACAGGAAGCCAAG 900
Qy 901 AGGCAGAGCATCAAGACAAAGTGGATGAGTTCTTGCTTATCAGGGCAGTCCGATGAT 960
Db 901 AGGCAGAGCATCAAGACAAAGTGGATGAGTTCTTGCTTATCAGGGCAGTCCGATGAT 960
Qy 961 GTCCCTCAGGGAAGGACATAGTTGACGCAAGGAAAAAGAACAAAGCAAACTTGGAACT 1020
Db 961 GTCCCTCAGGGAAGGACATAGTTGACGCAAGGAAAAAGAACAAAGCAAACTTGGAACT 1020
Qy 1021 GCGGATTGCTCTGTCTCTGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATCCAG 1080
Db 1021 GCGGATTGCTCTGTCTCTGGGACCCAGAAAGGACTGGAGACTGCAAAAGGACGATCCAG 1080
Qy 1081 GCAGCAGCTATAGCAGAGAAATAATATTCATCATGAGTTCTTTAAACTCTGGTTCA 1140
Db 1081 GCAGCAGCTATAGCAGAGAAATAATATTCATCATGAGTTCTTTAAACTCTGGTTCA 1140
Qy 1141 CCTAACAAAGCTGAAAACAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCACC 1200
Db 1141 CCTAACAAAGCTGAAAACAAAAGGACCCAGAAAGACACGGGTGCTGAAAAGTCAACCACC 1200
Qy 1201 ACTTCAGCTGACCTTAAGTCAGACAAAGCCAACTTTTATCATCTCCAGGAGACCCAAAGGGCT 1260
Db 1201 ACTTCAGCTGACCTTAAGTCAGACAAAGCCAACTTTTATCATCTCCAGGAGACCCAAAGGGCT 1260
Qy 1261 GGCNAGAAATTCNAAAGGATGCNACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAA 1320
Db 1261 GGCNAGAAATTCNAAAGGATGCNACCCATCGGGGCACACAGTCCGTGACAAACCCCTGAA 1320
Qy 1321 CCTCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db 1321 CCTCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Qy 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAAATGTGGTG 1440
Db 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAAATGTGGTG 1440
Qy 1441 TGTGAGTCACGAGTAGAGATTATAAGTCCNAGAAAGTAGNATCAGCTTACNACAGTG 1500
Db 1441 TGTGAGTCACGAGTAGAGATTATAAGTCCNAGAAAGTAGNATCAGCTTACNACAGTG 1500
Qy 1501 GACCTCAACGAAGAGATGCTGCACCTGAAACCCACAGAAAGGAACTCAAAAGGAGAA 1560
Db 1501 GACCTCAACGAAGAGATGCTGCACCTGAAACCCACAGAAAGGAACTCAAAAGGAGAA 1560
Qy 1561 AGCAAAACAAAGAACCTCTCTGATGGCGTTTCTCAGACAAATGTGAGTGAAGGGATGGA 1620
Db 1561 AGCAAAACAAAGAACCTCTCTGATGGCGTTTCTCAGACAAATGTGAGTGAAGGGATGGA 1620
Qy 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCC 1680
Db 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCC 1680
Qy 1681 ACAGAAAAGACTATCACACCGCCAGAGCCTGAAACCAACAGGAGCACCAAGAGGGTAAA 1740
Db 1681 ACAGAAAAGACTATCACACCGCCAGAGCCTGAAACCAACAGGAGCACCAAGAGGGTAAA 1740
Qy 1741 GAGGGCTCTCTGAAAGGACAAAGATCAGAGCCGAGATGAACAGAGAGAAAGCAAG 1800
Db 1741 GAGGGCTCTCTGAAAGGACAAAGATCAGAGCCGAGATGAACAGAGAGAAAGCAAG 1800
Qy 1801 CAGCAACCCAAAGAACAGCCCTGACACAGAGAGCCAGGTGGACCACTCACTG 1860
Db 1801 CAGCAACCCAAAGAACAGCCCTGACACAGAGAGCCAGGTGGACCACTCACTG 1860
Qy 1861 CAGAAATGGGACAAAGCTCCAAAAGAGACTTGAAGCGGACAGCTCCCTTTGGGGGCTTC 1920
Db 1861 CAGAAATGGGACAAAGCTCCAAAAGAGACTTGAAGCGGACAGCTCCCTTTGGGGGCTTC 1920
Qy 1921 TTTAAAGGCTGGGACCAAGCCGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCC 1980


```
Db 1921 TTTAAAGCGCTGGGACCAAGCGGATGTTGATGCTCAAGTGCAAAAGAGCCAGTATCC 1980
Qy 1981 ATCGGACAGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCACAGGTTCTCTG 2040
Db 1981 ATCGGACAGTTGGCAAAACCCAAAGTAAACAAATCAGCAGCGTTCCACAGGTTCTCTG 2040
Qy 2041 CCACCAAGATGTTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATTC 2100
Db 2041 CCACCAAGATGTTCTCTCTTACTCCATCTCTCTCCCAACACGCTCCATGTATATTC 2100
Qy 2101 TTCTGATGCCAGCAAAATCTGCTAGAAATTAAGCCCGAGCTGTTGTATATGA 2160
Db 2101 TTCTGATGCCAGCAAAATCTGCTAGAAATTAAGCCCGAGCTGTTGTATATGA 2160
Qy 2161 GGTGTATTTATTAAGTCTCTGTCAGTCTTTTCTGCAAAATAACAGTAAAGATGTTTA 2220
Db 2161 GGTGTATTTATTAAGTCTCTGTCAGTCTTTTCTGCAAAATAACAGTAAAGATGTTTA 2220
Qy 2221 GCAGGTCACTAGTTGGTCTAGAGAGTCGATGATCACAAGCAGGAAGGGAGGAATA 2280
Db 2221 GCAGGTCACTAGTTGGTCTAGAGAGTCGATGATCACAAGCAGGAAGGGAGGAATA 2280
Qy 2281 GAGGAATGTTTGGGTTTAAAGTGAATAATGGCAGTGTGGCGGGGCTGTGCTCTC 2340
Db 2281 GAGGAATGTTTGGGTTTAAAGTGAATAATGGCAGTGTGGCGGGGCTGTGCTCTC 2340
Qy 2341 GCCTGTAAATCTCAGCACCTTTGGGAGCCGAGGAGTGATCACCTGAGGTCAGGAGTTC 2400
Db 2341 GCCTGTAAATCTCAGCACCTTTGGGAGCCGAGGAGTGATCACCTGAGGTCAGGAGTTC 2400
Qy 2401 AAGACTAGCTGGCCAAACATCATGAACCCCGTCTCTACTAAATAACAAAAATTAGCCA 2460
Db 2401 AAGACTAGCTGGCCAAACATCATGAACCCCGTCTCTACTAAATAACAAAAATTAGCCA 2460
Qy 2461 GGCATGTTGGCACACACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGCT 2520
Db 2461 GGCATGTTGGCACACACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGCT 2520
Qy 2521 TGTATCCAGGAGGTGGAGGTTGCAAGTGGAGCGAAGTTGCACCTTGCACCTCCACCTGGG 2580
Db 2521 TGTATCCAGGAGGTGGAGGTTGCAAGTGGAGCGAAGTTGCACCTTGCACCTCCACCTGGG 2580
Qy 2581 CGACAGACAGATTTCTATCAAAAAAAGAGCAGTGGCAAGTAAATGATAGAGAGAAA 2640
Db 2581 CGACAGACAGATTTCTATCAAAAAAAGAGCAGTGGCAAGTAAATGATAGAGAGAAA 2640
Qy 2641 TGCTGCTAGAAGGAATTAAGCGTTTCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGAA 2700
Db 2641 TGCTGCTAGAAGGAATTAAGCGTTTCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGAA 2700
Qy 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGT 2760
Db 2701 GGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGT 2760
Qy 2761 TGGTGTGTTGCCAATTTCTAGACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAG 2820
Db 2761 TGGTGTGTTGCCAATTTCTAGACGGAATGTGTGGGTATAGAAAAAGGAATGAATAAG 2820
Qy 2821 CGTTGTTTTTCAATPAGGTCCTTGAAGTTATTGATCAGAGGGGAAAGATTGACTGGGG 2880
Db 2821 CGTTGTTTTTCAATPAGGTCCTTGAAGTTATTGATCAGAGGGGAAAGATTGACTGGGG 2880
Qy 2881 AGGGCTTAAATGATTGGGAAAAACAATTGCTTTTGGGCTCAGTGAACAACGGCAAGAT 2940
Db 2881 AGGGCTTAAATGATTGGGAAAAACAATTGCTTTTGGGCTCAGTGAACAACGGCAAGAT 2940
Qy 2941 TACAATTTAAAAAATAAAAAAATACTCGAGCTAGTTCTCTCTCTCTCTCTCTCTCTCGGCGA 3000
Db 2941 TACAATTTAAAAAATAAAAAAATACTCGAGCTAGTTCTCTCTCTCTCTCTCTCTCTCGGCGA 3000
Qy 3001 ATTCGATATCAAGCTTATCGATACCGTGAACCTCGAGGGGGGGCCGGTACCCCAATTCGC 3060
Db 3001 ATTCGATATCAAGCTTATCGATACCGTGAACCTCGAGGGGGGGCCGGTACCCCAATTCGC 3060
```

```
Db 3001 ATTCGATATCAAGCTTATCGATACCGTGAACCTCGAGGGGGGGCCGGTACCCCAATTCGC 3060
Qy 3061 CCTATA 3066
Db 3061 CCTATA 3066

RESULT 2
ADQ23347
ID ADQ23347 standard; DNA; 3418 BP.
XX AC ADQ23347;
XX AC AC
XX 26-AUG-2004 (first entry)
XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 6167.
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX ds.
XX Homo sapiens.
XX OS
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.
XX Example 2; SEQ ID NO 6167; 210pp; English.
XX The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
XX The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.
XX Sequence 3418 BP; 1029 A; 763 C; 842 G; 669 T; 0 U; 115 Other;

Query Match 92.3%; Score 2828.8; DB 12; Length 3418;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2833; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 127 GGACAGAGGCTCCACCGACAGCCAGGCACTGGGAGACGACCTGGAGACCCAGGACCT 186
Db 40 GGAGCCAGGCTCCACCGACAGCCAGGCACTGGGAGACGACCTGGAGACCCAGGACCT 99

Qy 187 GTGAGGAGCAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCA 246
Db 100 GTGAGGAGCAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCA 159

Qy 247 GGACCAATGGGTATACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGATGAACCA 306
Db 247 GGACCAATGGGTATACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGATGAACCA 306
```

Db 160 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAAATGAACCA 219
Qy 307 GAACAGAGACTTACCAAGGACCAACCGTCTGCTCTGAACGGGGTTCAGTGGTGTGTCG 366
Db 220 GAACAGAGACTTACCAAGGACCAACCGTCTGCTCTGAACGGGGTTCAGTGGTGTGTCG 279
Qy 367 ACCACACAGTTCAGCACTTAGAGAAAGTGCATTGGGAATAAGTGTCAAGA CGGATAAT 426
Db 280 ACCACACAGTTCAGCACTTAGAGAAAGTGCATTGGGAATAAGTGTCAAGA CGGATAAT 339
Qy 427 GTGGCCACTTCTCCCGGAGACAAACCGAGATAGTGTGTGGGATGCCAACGGAAAG 486
Db 340 GTGGCCACTTCTCCCGGAGACAAACCGAGATAGTGTGTGGGATGCCAACGGAAAG 399
Qy 487 AATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTTTTCTTGATG 546
Db 400 AATCTTGGGAAGAGGCCAAACCCGAGGACCAAGCTGTCTAAATCTCGTTTTTTCTTGATG 459
Qy 547 CTCTCTCGGCCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTATCCCTTTGGATCA 606
Db 460 CTCTCTCGGCCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTATCCCTTTGGATCA 519
Qy 607 GTGAAGCTTGATGTAGTCCATTAAGTCTCCAGCGAAACAAAGACCCCAAGTGAAGCTGG 666
Db 520 GTGAAGCTTGATGTAGTCCATTAAGTCTCCAGCGAAACAAAGACCCCAAGTGAAGCTGG 579
Qy 667 ACATCTCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCCAAGGSCACGCCCG 726
Db 580 ACATCTCGGTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCCAAGGSCACGCCCG 639
Qy 727 GCCCAAGACAAAGTCTCTCTCGCGCCAGGGATCCACGCTTCTCCACCTTGACACAGG 786
Db 640 GCCCAAGACAAAGTCTCTCTCGCGCCAGGGATCCACGCTTCTCCACCTTGACACAGG 699
Qy 787 GGAGCAGAGAGAGTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTC 846
Db 700 GGAGCAGAGAGAGTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTC 759
Qy 847 AAGCTGGACAGGACAGGAAAGGTGCCAGTGCACAGCCAAACAGGAGGCCAAGAGGCA 906
Db 760 AAGCTGGACAGGACAGGAAAGGTGCCAGTGCACAGCCAAACAGGAGGCCAAGAGGCA 819
Qy 907 GAGCATCAAGACAAAGTGGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCT 966
Db 820 GAGCATCAAGACAAAGTGGATGAGTTCCTGGCTTATCAGGCGAGTCCGATGATGTCCCT 879
Qy 967 GCAGGAAAGACATAGTTGACGGCAAGGAAAGAGAGACAAAGACTTGGAACTGCCGAT 1026
Db 880 GCAGGAAAGACATAGTTGACGGCAAGGAAAGAGAGACAAAGACTTGGAACTGCCGAT 939
Qy 1027 TGCTCTGCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCACAGGCAGCA 1086
Db 940 TGCTCTGCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCACAGGCAGCA 999
Qy 1087 GCTATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGTTTACCTTAAC 1146
Db 1000 GCTATAGCAGAGATAATAATTCATCATGATTTCTTTAAACTCTGTTTACCTTAAC 1059
Qy 1147 AAAGCTGAAACAAAAGGACCCAGAGACACGGGTGTGAAAAGTCAACCACTTCA 1206
Db 1060 AAAGCTGAAACAAAAGGACCCAGAGACACGGGTGTGAAAAGTCAACCACTTCA 1119
Qy 1207 GCTGACCTTAAGTCAGACAAAGCCAACTTATATCCAGGAGACCCAGGGGCTGGCAAG 1266
Db 1120 GCTGACCTTAAGTCAGACAAAGCCAACTTATATCCAGGAGACCCAGGGGCTGGCAAG 1179
Qy 1267 AATTCCAAAGGATGCAACCCCTCGGGGACACACAGTCCGTGACAAACCCCTGAACTGCG 1326
Db 1180 AATTCCAAAGGATGCAACCCCTCGGGGACACACAGTCCGTGACAAACCCCTGAACTGCG 1239
Qy 1327 AAGGAGGACCAAGGAGAAATCAGGACCGACTCTCTGGCTCTGGGCAAACTGTTTTGG 1386
Db 1240 AAGGAGGACCAAGGAGAAATCAGGACCGACTCTCTGGCTCTGGGCAAACTGTTTTGG 1299

Qy 1387 AAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGGTGTGAG 1446
Db 1300 AAAAAGTCAGTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGGTGTGAG 1359
Qy 1447 TCACCACTAGAGATTATAAGTCCAAGGAAGTGAATCAGCCTTACAAACAGTGGACCTC 1506
Db 1360 TCACCACTAGAGATTATAAGTCCAAGGAAGTGAATCAGCCTTACAAACAGTGGACCTC 1419
Qy 1507 AACGAAGGAGATCTGCACCTGACCCACAGNAGCGGAACTCAAAAGAGAAAGCAAA 1566
Db 1420 AACGAAGGAGATCTGCACCTGACCCACAGAGCGGAACTCAAAAGAGAAAGCAAA 1479
Qy 1567 CCAAGAAACCTCTCTGATGGGCTTTCTCAGACAAATGTCAAGTGAAGGGATGAGGGATC 1626
Db 1480 CCAAGAAACCTCTCTGATGGGCTTTCTCAGACAAATGTCAAGTGAAGGGATGAGGGATC 1539
Qy 1627 ACCCACTCAGAGAAATAATGGGAAAGATCTCAGCTGCCAAACATCAGACTCCACAGAA 1686
Db 1540 ACCCACTCAGAGAAATAATGGGAAAGATCTCAGCTGCCAAACATCAGACTCCACAGAA 1599
Qy 1687 AAGNACTATCACCGCCAGAGCCTGAACCAACAGGAGCACACAGAGAGGCTAAAGAGGC 1746
Db 1600 AAGNACTATCACCGCCAGAGCCTGAACCAACAGGAGCACACAGAGAGGCTAAAGAGGC 1659
Qy 1747 TCCTCGAAGGACAAAGAGTCAGCAGCGGAGATGAACCAAGCAGAGAGAGCAACAGCAGGAA 1806
Db 1660 TCCTCGAAGGACAAAGAGTCAGCAGCGGAGATGAACCAAGCAGAGAGAGCAACAGCAGGAA 1719
Qy 1807 GCCAAAGAACAGCCCAAGTGCACAGAGGCGCAACGCTGGACACGAATCTACTCAGAAAT 1866
Db 1720 GCCAAAGAACAGCCCAAGTGCACAGAGGCGCAACGCTGGACACGAATCTACTCAGAAAT 1779
Qy 1867 GGGGACAGCTCCAAAGAGAGCCTGAGAAGCGGAGAGTCCCTTGGGGGCTCTTTTAAA 1926
Db 1780 GGGGACAGCTCCAAAGAGAGCCTGAGAAGCGGAGAGTCCCTTGGGGGCTCTTTTAAA 1839
Qy 1927 GGCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1986
Db 1840 GGCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGA 1899
Qy 1987 CCAGTTGGGCAAAACCAAGTAAACAAATCAGACAGCGGTTCCCAACAGGTTCTCTGCCACCA 2046
Db 1900 CCAGTTGGGCAAAACCAAGTAAACAAATCAGACAGCGGTTCCCAACAGGTTCTCTGCCACCA 1959
Qy 2047 AGATGTGTTCTCTTACTCCATCTCTCCCAACACGCTCCATGTATATATTCTTCTGA 2106
Db 1960 AGATGTGTTCTCTTACTCCATCTCTCCCAACACGCTCCATGTATATATTCTTCTGA 2019
Qy 2107 TGGCCAGCAAAATGAAATTTGCTCTAGAAATTTAAGCCCGAGCTGTTGTATATTGAGGTGTA 2166
Db 2020 TGGCCAGCAAAATGAAATTTGCTCTAGAAATTTAAGCCCGAGCTGTTGTATATTGAGGTGTA 2079
Qy 2167 TTAATTAAGTCTCTGGTCCAGCTCTTTCTGCAAAATAACAGTAAAGATGGTTTAGCAGGT 2226
Db 2080 TTAATTAAGTCTCTGGTCCAGCTCTTTCTGCAAAATAACAGTAAAGATGGTTTAGCAGGT 2139
Qy 2227 CACCTAGTTGGTCAGAAAGAGTCGATCATCAACAGCAGGAAAGAGGGAGAAATAGAGAA 2286
Db 2140 CACCTAGTTGGTCAGAAAGAGTCGATCATCAACAGCAGGAAAGAGGGAGAAATAGAGAA 2199
Qy 2287 TGTCTTCCGGTTAAGTGATGAAATTTGTCAGTGTGGTGGCGGCGTGGTGGCTCTCGCTGT 2346
Db 2200 TGTCTTCCGGTTAAGTGATGAAATTTGTCAGTGTGGTGGCGGCGTGGTGGCTCTCGCTGT 2259
Qy 2347 AATCTCAGACTTTGGGAGGCCGAGGAGTGGATCACCTGAGTTCAGGATTCAGACT 2406
Db 2260 AATCTCAGACTTTGGGAGGCCGAGGAGTGGATCACCTGAGTTCAGGATTCAGACT 2319
Qy 2407 AGCCTGGCACAACATCATGAAACCCGCTCTCTACTTAAATAACAAAAATAGCCAGGCAATG 2466
Db 2320 AGCCTGGCACAACATCATGAAACCCGCTCTCTACTTAAATAACAAAAATAGCCAGGCAATG 2379

QY 2467 GTGGCACACCTGTAGTCCAGCTACTCGGAGGCCAACGACGAGAACCGCTTGTTACC 2526
DB 2380 GTGGCACACCTGTAGTCCAGCTACTCGGAGGCCAACGACGAGAACCGCTTGTTACC 2439
QY 2527 CAGGAGGTGGAGGTTCAGTGGAGCCGAAGTTGACCATTCACCTCCACCTGGGGACAG 2586
DB 2440 CAGGAGGTGGAGGTTCAGTGGAGCCGAAGTTGACCATTCACCTCCACCTGGGGACAG 2499
QY 2587 ACACAGATTCATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAGAGAAATGCTGC 2646
DB 2500 ACACAGATTCATCAAAAAAAGGCGAGTGCACAGTAAGTTATAGAGAGAAATGCTGC 2559
QY 2647 TAGAGGAATTAAGCGTTGTAGTAAACCGGTCTCATCTCTTAAGCTTTGAAGAAGGGAGA 2706
DB 2560 TAGAGGAATTAAGCGTTGTAGTAAACCGGTCTCATCTCTTAAGCTTTGAAGAAGGGAGA 2619
QY 2707 CGAAATCCATTTGTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTG 2766
DB 2620 CGAAATCCATTTGTTAAATTCACATCTCAAGAGGGAGAACCCGGGCTGTGTTGGGTG 2679
QY 2767 GTTGCCAAATTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGT 2826
DB 2680 GTTGCCAAATTCCTAGAACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGT 2739
QY 2827 TTTTCAAAATAGGCTCTTGTAGTTATTGATGAGAGGGAGAAAGATTGACTGGGAGGGCT 2886
DB 2740 TTTTCAAAATAGGCTCTTGTAGTTATTGATGAGAGGGAGAAAGATTGACTGGGAGGGCT 2799
QY 2887 TAAATGATTTGGGAAAAACAATTGCTTTTGGGCTCAGTGACCAACGGCAAGATTACAAAC 2946
DB 2800 TAAATGATTTGGGAAAAACAATTGCTTTTGGGCTCAGTGACCAACGGCAAGATTACAAAC 2859
QY 2947 TTAATAAAAAAAAAAAAAA 2967
DB 2860 TTAATAAAAAAAAAAAAAATAA 2880

RESULT 3
AAS81115
ID AAS81115 standard; cDNA; 2813 BP.
XX
AC AAS81115;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16919.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
XX
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
PI WPI; 2001-639362/73.
XX
DR P-FSDB; ASB16928.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX

PS Claim 1; SEQ ID NO 16919; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (II) is useful in gene therapy techniques to restore normal
CC activity of (III) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;
Query Match 91.7%; Score 2813; DB 5; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 136 CTCACACGACAGCAGGACCTGGGCGAGCAGCCACTGGAGACCAGGACCCCTGTGCAGGAG 195
DB 1 CTCACACGACAGCAGGACCTGGGCGAGCAGCCACTGGAGACCAGGACCCCTGTGCAGGAG 60
QY 196 CAGCTCCGGGTGCACACGAGGGGACTGAAAGATACTCCACAGGGGTCCAGTCAGGAGCAATG 255
DB 61 CAGCTCCGGGTGCACACGAGGGGACTGAAAGATACTCCACAGGGGTCCAGTCAGGAGCAATG 120
QY 256 GGTAAACAAATGAGTGTTCCTCAAGAGTTGAAGACCAAGAGAAATGAACCAAGACAGAG 315
DB 121 GGTAAACAAATGAGTGTTCCTCAAGAGTTGAAGACCAAGAGAAATGAACCAAGACAGAG 180
QY 316 ACTTACAGGACACGCGTCTGCTGAACGGGGTTCAGTGGTGGTCCGACCCACACA 375
DB 181 ACTTACAGGACACGCGTCTGCTGAACGGGGTTCAGTGGTGGTCCGACCCACACA 240
QY 376 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGACGGATAATGTGGCCACT 435
DB 241 GTTCAGCACTTAGAGGAAGTCGACTTGGGAATAGTGTCAAGACGGATAATGTGGCCACT 300
QY 436 TCTTCCCCCGAGACAAACGGAGATAAGTGTCTGTGGGATGCCAACCGGAAAGAAATCTTGGG 495
DB 301 TCTTCCCCCGAGACAAACGGAGATAAGTGTCTGTGGGATGCCAACCGGAAAGAAATCTTGGG 360
QY 496 AAAGAGGCCAAACCCGAGGACCCAGCTCTGCTAAATCTGTTTTTTTCTGATGCTCTCTGG 555
DB 361 AAAGAGGCCAAACCCGAGGACCCAGCTCTGCTAAATCTGTTTTTTTCTGATGCTCTCTGG 420
QY 556 CCTGTACCAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 615
DB 421 CCTGTACCAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTT 480
QY 616 GATGTCACTCCAATAAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTGGACACTTCG 675
DB 481 GATGTCACTCCAATAAAGCTCCAGCGAACCAAGACCCCAAGTGAGAGCTGGACACTTCG 540
QY 676 GTGCGAGCTGCACCGGGCAGACACAGATAAACCCTCCAGGCGACGCCCGCCCAAGAC 735
DB 541 GTGCGAGCTGCACCGGGCAGACACAGATAAACCCTCCAGGCGACGCCCGCCCAAGAC 600
QY 736 AAGTGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCCACTTGACAGAGGGGACAGGA 795
DB 601 AAGTGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCCACTTGACAGAGGGGACAGGA 660

RESULT 4	
AAD38092	
ID	AAD38092 standard; DNA; 2813 BP.
XX	
AC	AAD38092;
XX	
DT	10-SEP-2002 (first entry)
XX	
DE	Human BCAS1 (breast cancer amplified sequence 1) DNA.
XX	
KW	Human; BCAS1; breast cancer amplified sequence 1; AIBCL1; inflammation; amplified in breast cancer 1; NABCL1; novel amplified in breast cancer 1; hyperproliferative disorder; breast; prostate; cancer; prophylaxis; infection; antisense therapy; cytostatic; antiinflammatory; tumour; gene; ds.
KW	ds.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
FT	118..1872
FT	/*tag= a
FT	/product= "Human BCAS1 protein"
XX	
PN	WO200231136-A1.
XX	
PD	18-APR-2002.
XX	
PP	09-OCT-2001; 2001WO-US031484.
XX	
PR	11-OCT-2000; 2000US-00689255.
XX	
PA	(ISIS-) ISIS PHARM INC.
XX	
PI	Cowser LM, Freier SM;
XX	
DR	P-PSDB; AAE23791.
XX	
PT	New antisense compounds targeted to a nucleic acid molecule encoding BCAS1, useful for treating diseases or conditions associated with BCAS1, such as hyperproliferative disease, particularly breast or prostate cancer.
XX	
PS	Example 10; Page 92-95; 104pp; English.
XX	
CC	The invention relates to antisense compounds, compositions and methods for modulating the expression of BCAS1 (breast cancer amplified sequence 1, also known as AIBCL1 for amplified in breast cancer 1 and NABCL1 for novel amplified in breast cancer 1). The antisense compounds of the invention are useful for treating an animal having a disease or condition associated with BCAS1, such as hyperproliferative disorders including breast or prostate cancer. These compounds are also used as research reagents and diagnostics; to distinguish between functions of various members of a biological pathway; in the treatment of a disease or disorder, which can be treated by modulating the expression of BCAS1; as prophylaxis, e.g. to prevent or delay infection, inflammation or tumour formation; and as probes or primers. These antisense compounds are used in antisense therapy. The present sequence is human BCAS1 DNA
XX	
SQ	Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;
XX	
Query Match	91.7%; Score 2813; DB 6; Length 2813;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 2813; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
XX	
Qy	136 CTCCACCCAGCCAGGCACTGGGCGACCGCACTGGAGCCAGGACCTGTGCAGGAG 195
Db	1 CTCCACCCAGCCAGGCACTGGGCGACCGCACTGGAGCCAGGACCTGTGCAGGAG 60
XX	
Qy	196 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 255
Db	61 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 120

Qy	256 GGTACCAAAATGAGTGTTCCTCCCAAGAGCTTGAAGACCAAGAGAAATGAACCAAGCAGAG 315
Db	121 GGTACCAAAATGAGTGTTCCTCCCAAGAGCTTGAAGACCAAGAGAAATGAACCAAGCAGAG 180
Qy	316 ACTTACCAAGGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGTCGACCCACACA 375
Db	181 ACTTACCAAGGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGTCGACCCACACA 240
Qy	376 GTTCAGCACTTAGAGGAAGTGCATTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
Db	241 GTTCAGCACTTAGAGGAAGTGCATTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
Qy	436 TCTTCCCGGAGACAAACGCGATAGTCTGTTCGGATGCCAACGGAAAGAAATCTTGGG 495
Db	301 TCTTCCCGGAGACAAACGCGATAGTCTGTTCGGATGCCAACGGAAAGAAATCTTGGG 360
Qy	496 AAAGAGGCCAAACCCGAGGCACAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCGG 555
Db	361 AAAGAGGCCAAACCCGAGGCACAGCTGCTAAATCTCGTTTTTTTCTTGATGCTCTCGG 420
Qy	556 CTTGTACAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTTGGATTCAGTGAAGTT 615
Db	421 CTTGTACAGGACGTACCGGAGACCAAGCCGAGATTTCATCCCTTTGGATTCAGTGAAGTT 480
Qy	616 GATGTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTGAGAGCTGGACATTCG 675
Db	481 GATGTCAGCTCCAATAAAGCTCCAGCGAAACAAAGACCCAAAGTGAGAGCTGGACATTCG 540
Qy	676 GTGGCAGCTGGACCGGGGACGACACAGATAAAACCCAGGGCAGCCGCCGCCCAAGAC 735
Db	541 GTGGCAGCTGGACCGGGGACGACACAGATAAAACCCAGGGCAGCCGCCGCCCAAGAC 600
Qy	736 AAGTCTCTCTGCGCGCAGGGATCCCAACGCTTCTCCACCTGACAGGGGGAGCAGGA 795
Db	601 AAGTCTCTCTGCGCGCAGGGATCCCAACGCTTCTCCACCTGACAGGGGGAGCAGGA 660
Qy	796 GGAGAGCTCCCTCCNAGCCCAAGACTCCAGCTTTTTTTTGACAAATTTCTCAGCTGAC 855
Db	661 GGAGAGCTCCCTCCNAGCCCAAGACTCCAGCTTTTTTTTGACAAATTTCTCAGCTGAC 720
Qy	856 AAGGACAGGAAAAGGTGCCAGGTGACAGCAACAGAGGCAAGCAAGGGCAGAGCATCAA 915
Db	721 AAGGACAGGAAAAGGTGCCAGGTGACAGCAACAGAGGCAAGCAAGGGCAGAGCATCAA 780
Qy	916 GACAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTGACGGGAAG 975
Db	781 GACAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTGACGGGAAG 840
Qy	976 GACATAGTTGAGCGCAAGGAAAAGAGGACCAAGAACTTGGAACTGCGGATGCTCTGTC 1035
Db	841 GACATAGTTGAGCGCAAGGAAAAGAGGACCAAGAACTTGGAACTGCGGATGCTCTGTC 900
Qy	1036 CCTGGGACCCAGAGGACTGAGACTGCAAGGACGATTCAGGAGCAGAGCTATAGCA 1095
Db	901 CCTGGGACCCAGAGGACTGAGACTGCAAGGACGATTCAGGAGCAGAGCTATAGCA 960
Qy	1096 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1155
Db	961 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1020
Qy	1156 ACAAAAAGGACCCAGAGACACCGGTGCTGAAAAGTCAACCACTTCAGCTGACCTT 1215
Db	1021 ACAAAAAGGACCCAGAGACACCGGTGCTGAAAAGTCAACCACTTCAGCTGACCTT 1080
Qy	1216 AAGTCAGCAAGCAACTTTTACATCCAGAGACCCAGGGGCTGGCAAGAAATTTCCAAA 1275
Db	1081 AAGTCAGCAAGCAACTTTTACATCCAGAGACCCAGGGGCTGGCAAGAAATTTCCAAA 1140
Qy	1276 GGATCCAAACCATCCGGGACACACAGTCCGCTGCAACCCCTGAACTTCGCAAGGAAGGC 1335
Db	1141 GGATCCAAACCATCCGGGACACACAGTCCGCTGCAACCCCTGAACTTCGCAAGGAAGGC 1200
Qy	1336 ACCAAGGAGAAATCAGGACCCCACTCTCTGCTCTGGGCAAACTGTTTTTGGAAAAAGTCA 1395

The present invention describes a method of testing for bronchial asthma or chronic obstructive pulmonary disease. The method comprises determining the expression level of a marker gene in a biological sample from a subject, comparing the expression level determined with the expression level of the marker gene in a biological sample from a healthy subject, and judging whether the subject has bronchial asthma or chronic obstructive pulmonary disease. The marker gene comprises: (a) a group of genes (S1) whose expression levels increase when respiratory epithelial cells are stimulated with interleukin-13; or (b) a group of genes (S2) whose expression levels decrease when respiratory epithelial cells are stimulated with interleukin-13. Also described: (1) a reagent (I) for testing for bronchial asthma or chronic obstructive pulmonary disease; (2) a kit for screening for a candidate compound for a therapeutic agent to treat bronchial asthma or chronic obstructive pulmonary disease; (3) an animal model for bronchial asthma or chronic obstructive pulmonary disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a method for producing an animal model for bronchial asthma or chronic obstructive pulmonary disease; (6) a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease, comprising the compound, a marker gene or an antisense nucleic acid corresponding to a portion of the marker gene, a ribozyme, a polynucleotide that suppresses the expression of the gene through an RNAi effect or an antibody recognising a protein encoded by a marker gene, and (7) a DNA chip for testing for bronchial asthma or a chronic obstructive pulmonary disease, on which a probe has been immobilised to assay a marker gene. (I) has respiratory and antilasthmatic activities, and can be used in gene therapy. The method is useful for testing for or screening for a therapeutic agent for bronchial asthma or chronic obstructive pulmonary disease. The present sequence is used in the exemplification of the present invention.

Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;

Query Match	91.7%;	Score 2813;	DB 12;	Length 2813;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2813;	Conservative	0;	Mismatches	0; Indels 0; Gaps 0
QY	136	CTCCACCGACGACGAGGACCTGGGCAGACGACGACCTGGAGACCCAGGACCCCTGTGCGAGGAG	195	
DB	1	CTCCACCGACGACGAGGACCTGGGCAGACGACGACCTGGAGACCCAGGACCCCTGTGCGAGGAG	60	
QY	196	CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACACAGGGGCTCAGCAGGAGCAATG	255	
DB	61	CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCACACAGGGGCTCAGCAGGAGCAATG	120	
QY	256	GGTAACCAATCAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGAG	315	
DB	121	GGTAACCAATCAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCAAGAGCAGAG	180	
QY	316	ACTTACCGAGCAACCGCTCTGCTCTGAACCGGGTTCCAGTGGTGGTGTCGACCCACACAC	375	
DB	181	ACTTACCGAGCAACCGCTCTGCTCTGAACCGGGTTCCAGTGGTGGTGTCGACCCACACAC	240	
QY	376	GTTTCAGCACTTACGAGGAAGTGCAGTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT	435	
DB	241	GTTTCAGCACTTACGAGGAAGTGCAGTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT	300	
QY	436	TCTTCCCGCGACACACGGAGTAAGTCTCTTGGGATGCCAACCGAAGAGATCTTGGG	495	
DB	301	TCTTCCCGCGACACACGGAGTAAGTCTCTTGGGATGCCAACCGAAGAGATCTTGGG	360	
QY	496	AAAGAGGCCAAACCCGAGCACCAGCTCTCTAAATCTCGTTTTTCTTGATGCTCTCTCGG	555	
DB	361	AAAGAGGCCAAACCCGAGCACCAGCTCTCTAAATCTCGTTTTTCTTGATGCTCTCTCGG	420	
QY	556	CCTGTACCAAGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT	615	
DB	421	CCTGTACCAAGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCAGTGAAGCTT	480	
QY	616	GATGTCAGCTCCAAATAAGCTTCCAGCAACCAAGACCCAGTGAGAGCTCGACACTTCGG	675	
DB	481	GATGTCAGCTCCAAATAAGCTTCCAGCAACCAAGACCCAGTGAGAGCTCGACACTTCGG	540	
QY	676	GTGGCAGCTGGACCGGGGCGAGACACAGATAAAACCCGAGGCAACGCCCGGCCCAAGAC	735	

Db 1621 GACAAGAAGTCAGCAGCCGAGATGAAACAAGCAGAGAGCAACAAGCAGGAGCCAAAGAA 1680
Qy 1816 CCAGCCAGTSCACAGAGCAGCCACGCTGTGACACGAACTCAGTCAGAAATGGGGACAAG 1875
Db 1681 CCAGCCAGTSCACAGAGCAGCCACGCTGTGACACGAACTCAGTCAGAAATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGAGAAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGAGAAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCATCGGACCAAGTTGGC 1860
Qy 1996 AARCCCAAGTAAACAATCAGCAGCGTTCCCAACAGGTTCTCCGCCCACCAAGATGTT 2055
Db 1861 AARCCCAAGTAAACAATCAGCAGCGTTCCCAACAGGTTCTCCGCCCACCAAGATGTT 1920
Qy 2056 CTCCTTACTCCATCTCCTCCCAACACGCTCCATGTATATATTTCTGTATGCCACGA 2115
Db 1921 CTCCTTACTCCATCTCCTCCCAACACGCTCCATGTATATATTTCTGTATGCCACGA 1980
Qy 2116 AATGAATTTGCTCTAGAAATTAAGCCGAGCTGTTGTATATTAGGTGTATTTAAG 2175
Db 1981 AATGAATTTGCTCTAGAAATTAAGCCGAGCTGTTGTATATTAGGTGTATTTAAG 2040
Qy 2176 TCTCTGTCAGTCTTTCTGCAAAATAACAGTAAAGATGTTTACAGGTCACCTAGTT 2235
Db 2041 TCTCTGTCAGTCTTTCTGCAAAATAACAGTAAAGATGTTTACAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAAGAGTCGATGATCACAAGCAGGAGGAGGGAATAGAGGAATGTTTCGG 2295
Db 2101 GGGTCAGAAGAGTCGATGATCACAAGCAGGAGGAGGGAATAGAGGAATGTTTCGG 2160
Qy 2296 GTTAAGTATGAATAATGTCAGTGTGCGCGGCGTGTGCTCTCGCTGTATCTCAGC 2355
Db 2161 GTTAAGTATGAATAATGTCAGTGTGCGCGGCGTGTGCTCTCGCTGTATCTCAGC 2220
Qy 2356 ACTTTGGAGCCGAGGAGGAGTGCATCAGTCAGTCAGGAGTTCAGACTAGCTCGCC 2415
Db 2221 ACTTTGGAGCCGAGGAGGAGTGCATCAGTCAGTCAGGAGTTCAGACTAGCTCGCC 2280
Qy 2416 AACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGTTGGCACAC 2475
Db 2281 AACATCATGAACCCCGTCTCTACTAAAAATACAAAAATAGCCAGGATGTTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGTTGTACCCAGGAGTG 2535
Db 2341 ACCTGTAGTCCAGTACTCGGGAGCCCAACGACGAGAACCGTTGTACCCAGGAGTG 2400
Qy 2536 GAGGTTGAGTGAGCCGAGGTTGACCATTTGCATCTCCACCTGGGCGACAGCAAGATT 2595
Db 2401 GAGGTTGAGTGAGCCGAGGTTGACCATTTGCATCTCCACCTGGGCGACAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGGAGTGAAGTATAGAGGAATGCTGTAGAAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGAAGTATAGAGGAATGCTGTAGAAGGAA 2520
Qy 2656 TTAAGCGTTGATGAACCGTGTCTATCTCTAAGCTTGAAGAGGAGACGAAATCC 2715
Db 2521 TTAAGCGTTGATGAACCGTGTCTATCTCTAAGCTTGAAGAGGAGACGAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGTGTGTTCCCAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGTGTGTTCCCAAT 2640
Qy 2776 TTCTTAGAACGGAATGTTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAAT 2835
Db 2641 TTCTTAGAACGGAATGTTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTCAAT 2700
Qy 2836 AGGTCCTTGAAGTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATGAT 2895
Db 2701 AGGTCCTTGAAGTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATGAT 2760

Qy 2896 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAACTT 2948
Db 2761 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACGGCAAGATTACAACTT 2813

RESULT 6
ADQ18993

ID ADQ18993 standard; DNA; 2813 BP.

XX ADQ18993;

DT 26-AUG-2004 (first entry)

XX Human soft tissue sarcoma-upregulated DNA - SEQ ID 1812.

XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.

OS Homo sapiens.

XX W02004048938-A2.

XX 10-JUN-2004.

XX 26-NOV-2003; 2003WO-US038193.

XX 26-NOV-2002; 2002US-0429739P.

XX (PROT-) PROTEIN DESIGN LABS INC.

XX Aziz N, Ginsburg WM, Zlotnik A;

XX WPI; 2004-441208/41.

XX Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
and comparing the gene expression, also useful in treating soft tissue
sarcoma.

Example 2; SEQ ID NO 1812; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
which comprises obtaining a first soft tissue sample from an individual
and a normal soft tissue sample from the same or different individual,
determining the expression of a gene in both samples and comparing the
expression of the gene in both soft tissue samples, where a higher level
of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has
cytostatic applications and may be useful for detecting soft tissue
sarcoma, possibly via gene therapy or vaccine production. The nucleic
acid sequences may be useful in diagnostic and screening applications.
The current sequence is that of a human soft tissue sarcoma-upregulated
DNA of the invention. The current sequence is not shown within the
specification per se but was submitted in CD format by the inventor.

Sequence 2813 BP; 868 A; 671 C; 758 G; 516 T; 0 U; 0 Other;

Query Match 91.7%; Score 2813; DB 12; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 136 CTCACCCAGCAGCCAGGAGCTGGGAGCAGCAGCTGAGAGCCAGGACCCCTGTGCAGGAG 195
Db 1 CTCACCCAGCAGCCAGGAGCTGGGAGCAGCAGCTGAGAGCCAGGACCCCTGTGCAGGAG 60
Qy 196 CAGTCCCGGTGACACAGGGGAGCTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 255
Db 61 CAGTCCCGGTGACACAGGGGAGCTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 120
Qy 256 GGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAGACCAAGAGATGACCAAGAGCCAGAG 315
Db 121 GGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAGACCAAGAGATGACCAAGAGCCAGAG 180

QY 316 ACTTACAGGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGAGTCCACACACA 375
DB |||||
DB 181 ACTTACAGGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTGGTGTGAGTCCACACACA 240
QY 376 GTTCAAGCACTAGAGGAAAGTGCAGCTTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
DB |||||
DB 241 GTTCAAGCACTAGAGGAAAGTGCAGCTTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
QY 436 TCTTCCCGGAGACAAACGGAGATAAGTGTGTGCGGATGCGAACCGGAAAGAAATCTTGGG 495
DB |||||
DB 301 TCTTCCCGGAGACAAACGGAGATAAGTGTGTGCGGATGCGAACCGGAAAGAAATCTTGGG 360
QY 496 AAGAGGCCAAACCCGAGGCCACCTGCTGAATCTCGTCTTTTCTTCTGATGCTCTCTCGG 555
DB |||||
DB 361 AAGAGGCCAAACCCGAGGCCACCTGCTGAATCTCGTCTTTTCTTCTGATGCTCTCTCGG 420
QY 556 CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTCTATCCCTTGGATCAGTGAAGCTT 615
DB |||||
DB 421 CCTGTACAGGACGTACCGGAGACCAAGCCGAGATTCTATCCCTTGGATCAGTGAAGCTT 480
QY 616 GATGTCAAGTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAGCTGACACTTCCG 675
DB |||||
DB 481 GATGTCAAGTCCAAATAAGTCTCAGCGAAACAAAGACCCAAAGTGAGCTGACACTTCCG 540
QY 676 GTGGCAGCTGACCCGGGAGGACACAGATAAACCCGAGGACCGCCCGGCCCAAGAC 735
DB |||||
DB 541 GTGGCAGCTGACCCGGGAGGACACAGATAAACCCGAGGACCGCCCGGCCCAAGAC 600
QY 736 AAGGTCTCTCTGCGCGCCAGGATCCACAGCTTCTCCACCTGAGACAGGGGGAGCAGGA 795
DB |||||
DB 601 AAGGTCTCTCTGCGCGCCAGGATCCACAGCTTCTCCACCTGAGACAGGGGGAGCAGGA 660
QY 796 GGAGAAAGTCCCTTCAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGAC 855
DB |||||
DB 661 GGAGAAAGTCCCTTCAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGAC 720
QY 856 AAGGACAGGAAAGGTCGAGTGACAGCCAAAGGAAAGCAGAGGGGAGAGCATCAA 915
DB |||||
DB 721 AAGGACAGGAAAGGTCGAGTGACAGCCAAAGGAAAGCAGAGGGGAGAGCATCAA 780
QY 916 GACAAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAG 975
DB |||||
DB 781 GACAAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGAG 840
QY 976 GACATAGTTGACGCAAGGAAAGGAGGACAAAGAACTTGGAACTGCGGATTTGCTCTGTC 1035
DB |||||
DB 841 GACATAGTTGACGCAAGGAAAGGAGGACAAAGAACTTGGAACTGCGGATTTGCTCTGTC 900
QY 1036 CTTGGGACCCAGAAAGGACTGAGACTGCAAGGACGATTCGCGGACAGCAGCTATAGCA 1095
DB |||||
DB 901 CTTGGGACCCAGAAAGGACTGAGACTGCAAGGACGATTCGCGGACAGCAGCTATAGCA 960
QY 1096 GAGATATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1155
DB |||||
DB 961 GAGATATAATTCATCATGATTTCTTTAAACTCTGTTTCACTTAAACAAAGCTGAA 1020
QY 1156 ACAAAAAGGACCCAGAGACACGGGTCTGAAAAGTCAACCACTTCAAGCTGACCTT 1215
DB |||||
DB 1021 ACAAAAAGGACCCAGAGACACGGGTCTGAAAAGTCAACCACTTCAAGCTGACCTT 1080
QY 1216 AAGTCAGACAAAGCCAACTTTATCCAGAGACCCAAAGGGGTGGCAAGAAATTCACAA 1275
DB |||||
DB 1081 AAGTCAGACAAAGCCAACTTTATACATCCAGAGACCCAAAGGGGTGGCAAGAAATTCACAA 1140
QY 1276 GGATGCAACCATCGGGGACACAGCTCCGTGACAACTTGAACCTGACCTGCGAAGAGGC 1335
DB |||||
DB 1141 GGATGCAACCATCGGGGACACAGCTCCGTGACAACTTGAACCTGCGAAGAGAGGC 1200
QY 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1395
DB |||||
DB 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAGTCA 1260

QY 1396 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGAGAAATGTGTTGTGAGTCCACAGTA 1455
DB |||||
DB 1261 GTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAAATGTGTTGTGAGTCCACAGTA 1320
QY 1456 GAGATTATAAGTCCAAAGAAAGTAAATCAGCTTTACAAACAGTGGACCTCAACGAAGGA 1515
DB |||||
DB 1321 GAGATTATAAGTCCAAAGAAAGTAAATCAGCTTTACAAACAGTGGACCTCAACGAAGGA 1380
QY 1516 GATGTGCACCTGAACCCACAGAACGCGAAATCTCAAAAGAGAGAAAGCAAAACCAAGAAC 1575
DB |||||
DB 1381 GATGTGCACCTGAACCCACAGAACGCGAAATCTCAAAAGAGAGAAAGCAAAACCAAGAAC 1440
QY 1576 TCTCTGATGGCGTTTCTCAGACAAATGTCACTGAAAGGGGATGAGGGATCACCCACTCA 1635
DB |||||
DB 1441 TCTCTGATGGCGTTTCTCAGACAAATGTCACTGAAAGGGGATGAGGGATCACCCACTCA 1500
QY 1636 GAAGAAATAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAACTATC 1695
DB |||||
DB 1501 GAAGAAATAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGAACTATC 1560
QY 1696 ACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAAGGGCTCTCTCGAAG 1755
DB |||||
DB 1561 ACACCGCAGAGCTGAAACCAACAGGAGCACCAAGAGGGTAAAGAGGGCTCTCTCGAAG 1620
QY 1756 GACAAGAGTCAAGCAGCGGAGATGAAACAGAGAGAGCAACAGCAGGAAAGCCAAAGAA 1815
DB |||||
DB 1621 GACAAGAGTCAAGCAGCGGAGATGAAACAGAGAGAGCAACAGCAGGAAAGCCAAAGAA 1680
QY 1816 CCAGCCAGTGCACAGAGCAGGCCACGCTGACACGAACTCACTGCAAAATGGGAGCAAG 1875
DB |||||
DB 1681 CCAGCCAGTGCACAGAGCAGGCCACGCTGACACGAACTCACTGCAAAATGGGAGCAAG 1740
QY 1876 CTCCAAAGAGACTGTAAGAGCGGACGAGCTCCCTTGGGGCTTTCTTTAAAGGCTTGGGA 1935
DB |||||
DB 1741 CTCCAAAGAGACTGTAAGAGCGGACGAGCTCCCTTGGGGCTTTCTTTAAAGGCTTGGGA 1800
QY 1936 CCAAGCGEATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCATCGAACAGTGGC 1995
DB |||||
DB 1801 CCAAGCGEATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCATCGAACAGTGGC 1860
QY 1996 AAACCCAAAGTAAACAAATCAGACGCTTCCACACAGGTTCTCTGCGCAACCAAGATGTT 2055
DB |||||
DB 1861 AAACCCAAAGTAAACAAATCAGACGCTTCCACACAGGTTCTCTGCGCAACCAAGATGTT 1920
QY 2056 CTCTTACTCCATCTCCTCCCAACACGCTTCCATATATATATATTTCTTCTGATGGCCAGCA 2115
DB |||||
DB 1921 CTCTTACTCCATCTCCTCCCAACACGCTTCCATATATATATATTTCTTCTGATGGCCAGCA 1980
QY 2116 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATTTGAGGTGATATTTTACG 2175
DB |||||
DB 1981 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTTGTATATTTGAGGTGATATTTTACG 2040
QY 2176 TCTCTGTGCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2235
DB |||||
DB 2041 TCTCTGTGCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2100
QY 2236 GGGTCAGAGAGTGCATGATCAACCAAGCAGGAAAGGAGGGAATAGAGGAATGTGTTCCG 2295
DB |||||
DB 2101 GGGTCAGAGAGTGCATGATCAACCAAGCAGGAAAGGAGGGAATAGAGGAATGTGTTCCG 2160
QY 2296 GTTAAAGTGAATAATGGCAGTGGTGGCGGGCTGTGGCTCTCGCTGTGTAATCTCAGC 2355
DB |||||
DB 2161 GTTAAAGTGAATAATGGCAGTGGTGGCGGGCTGTGGCTCTCGCTGTGTAATCTCAGC 2220
QY 2356 ACTTTGGAGGCGCAGGAGTGGATCACTGAGTCAAGGATTCAGAGCTTAGCTGGCC 2415
DB |||||
DB 2221 ACTTTGGAGGCGCAGGAGTGGATCACTGAGTCAAGGATTCAGAGCTTAGCTGGCC 2280
QY 2416 AACATCATGAAACCCCTCTCTACTAAATAATACAAAAATTAGCCAGGAGTGGTGGCACAC 2475
DB |||||
DB 2281 AACATCATGAAACCCCTCTCTACTAAATAATACAAAAATTAGCCAGGAGTGGTGGCACAC 2340
QY 2476 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGCAACGAGAAACCGCTTGTACCCAGGAGGTG 2535

Db 901 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGACGATTCCTCAGGACGAGCTATAGCA 960
Qy 1096 GAGATAATATTCATCATGATGTTCTTTAAACCTCTGGTTTACCTAAACAAGCTGAA 1155
Db 961 GAGATAATATTCATCATGATGTTCTTTAAACCTCTGGTTTACCTAAACAAGCTGAA 1020
Qy 1156 ACAAAAAGGACCCAGAGAGACGCGGTGCTGAAAAGTCAACCACACCTTCAGCTGACCTT 1215
Db 1021 ACNAAAAGGACCCAGAGAGACGCGGTGCTGAAAAGTCAACCACACCTTCAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGACCCCAAGGGCTGGCAAGATTTCCAAA 1140
Qy 1276 GGATGCACCATCGGGGACACACAGTCCGTGCAACCCCTGAACTTCGCAAGGAGGC 1335
Db 1141 GGATGCACCATCGGGGACACACAGTCCGTGCAACCCCTGAACTTCGCAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAGTCA 1260
Qy 1396 GTTAAAGAGACTCAGTCCCAAGGTCGAGGAGAAATGTGGTGTGTGAGTCAACAGTA 1455
Db 1261 GTTAAAGAGACTCAGTCCCAAGGTCGAGGAGAAATGTGGTGTGTGAGTCAACAGTA 1320
Qy 1456 GAGATTATAAGTCAAGAGAGTAGAATCAGCTTTAACAAGTGGAGCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCAAGAGAGTAGAATCAGCTTTAACAAGTGGAGCTCAACGAAGGA 1380
Qy 1516 GATGCTGACCTGAAACCCACAGAGCGGAACTCTCAAAAGAGAGAAAGCAACCAAGGACC 1575
Db 1381 GATGCTGACCTGAAACCCACAGAGCGGAACTCTCAAAAGAGAGAAAGCAACCAAGGACC 1440
Qy 1576 TCTCTGTGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGGAGGATCACCCACTCA 1635
Db 1441 TCTCTGTGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGGAGGATCACCCACTCA 1500
Qy 1636 GAAGAAATATAAGTGAAGAGACTCCAGCTGCGCAAAATCATGATCTCACAAGAAAGCTATC 1695
Db 1501 GAAGAAATATAAGTGAAGAGACTCCAGCTGCGCAAAATCATGATCTCACAAGAAAGCTATC 1560
Qy 1696 ACACCGCAGAGCTGAAACCAACAGGAGCACCACAGAGGGTAAAGGGGCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAAACCAACAGGAGCACCACAGAGGGTAAAGGGGCTCTCGAAG 1620
Qy 1756 GACAAGAGTCAGCAGCGCGAGATGAAACAGCAGAGAGCAACAGAGGAGGAGGAGGAGGAGG 1815
Db 1621 GACAAGAGTCAGCAGCGCGAGATGAAACAGCAGAGAGCAACAGAGGAGGAGGAGGAGGAGG 1680
Qy 1816 CCAGCCAGTCACAGAGCAGGCCACGCTGGACACAGAACTCACTGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTCACAGAGCAGGCCACGCTGGACACAGAACTCACTGCAAGATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGGAAGCGGACAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGGAAGCGGACAGTCCCTTGGGGCTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCACCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCACCAAGAGAGAGAGGAGGAGGAGGAGGAGGAGG 1860
Qy 1996 AAAACCAAGTAAACAAATCAGACCGGTTCCACAGGTTCTCTGCCACCAAGATGTTT 2055
Db 1861 AAAACCAAGTAAACAAATCAGACCGGTTCCACAGGTTCTCTGCCACCAAGATGTTT 1920
Qy 2056 CTCCTTACTCCTCTCTCCCAACAGCTCCATGATATATTTCTTCTGATGGCCAGCA 2115
Db 1921 CTCCTTACTCCTCTCTCCCAACAGCTCCATGATATATTTCTTCTGATGGCCAGCA 1980
Qy 2116 AATGAAATTCGCTAGAAAATTAAGCCCGAGCTGTGTATATTCAGGTGATATTTACG 2175
Db 1981 AATGAAATTCGCTAGAAAATTAAGCCCGAGCTGTGTATATTCAGGTGATATTTACG 2040

Qy 2176 TCTCTGGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTTAGCAGGTCACCTAGTT 2235
Db 2041 TCTCTGGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTTAGCAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAGAGTCGATGATCACCAGCAGGAAAGGGAGGAAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAGAGAGTCGATGATCACCAGCAGGAAAGGGAGGAAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCGCTGTGGCTCTCGCTCTGAATCTCAGC 2355
Db 2161 GTTAAAGTGAATAAATGGCAGTGGTGGCCGGCGCTGTGGCTCTCGCTCTGAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCAGGACAGGTGATCACTGAGGTCAAGGATTCAGACACTAGCCTGGCC 2415
Db 2221 ACTTTGGAGGCCAGGACAGGTGATCACTGAGGTCAAGGATTCAGACACTAGCCTGGCC 2280
Qy 2416 AACATCATGAAACCCCTCTCTACTAAATAATACAAAATATAGCCAGGATGTTGGCACAC 2475
Db 2281 AACATCATGAAACCCCTCTCTACTAAATAATACAAAATATAGCCAGGATGTTGGCACAC 2340
Qy 2476 ACTGTAGTCCAGCTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2535
Db 2341 ACTGTAGTCCAGCTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2400
Qy 2536 GAGTTGAGTGCAGCCGAAAGTTGCACCTTGCACCTCCACCTGGGCGACAGCAAGATT 2595
Db 2401 GAGTTGAGTGCAGCCGAAAGTTGCACCTTGCACCTCCACCTGGGCGACAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGAGGAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGCTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGCTAGAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGAGCAAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACGCGTCTCATCTCTTAAGCTTGAAGAGGAGAGCAAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAGGAGGAGAACCCGGGCTGTGTTGGTGGTTCCTCAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAGGAGGAGAACCCGGGCTGTGTTGGTGGTTCCTCAT 2640
Qy 2776 TTCTCAGAACGGAATGTGGGTATAGAAAAGGAATGAATAGCGTTGTTTTCAAAT 2835
Db 2641 TTCTCAGAACGGAATGTGGGTATAGAAAAGGAATGAATAGCGTTGTTTTCAAAT 2700
Qy 2836 AGGTCTCTTGTAAAGTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATCAT 2895
Db 2701 AGGTCTCTTGTAAAGTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATCAT 2760
Qy 2896 TTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAGATTCACCTT 2948
Db 2761 TTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGGCAAGATTCACCTT 2813

RESULT 8

ACC90605

ID ACC90605 standard; DNA; 2020 BP.

XX ACC90605;

AC ACC90605;

DX 12-AUG-2003 (first entry)

XX Human CGDP-28 encoding DNA.

DE Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX Human; cytosolic; antiarteriosclerotic; anticonvulsant; nootropic;

XX	03-APR-2003.	Db	126	GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAAATGAACCA	185
XX		Qy	307	GAAGCAGAGACTTTACCAGGACAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGTGCG	366
XX	26-SEP-2002; 2002WO-US031095.	Db	186	GAAGCAGAGACTTTACCAGGACAACCGCTCTGCTCTGAACGGGGTTCCAGTGGTGTGCG	245
XX	28-SEP-2001; 2001US-0326389P.	Qy	367	ACCCACACAGTTTCCAGCACTTAGAGGAAGTCCAGTCTGGGAATAAGTGTCAAGACGGATAAT	426
XX	05-OCT-2001; 2001US-0327380P.	Db	246	ACCCACACAGTTTCCAGCACTTAGAGGAAGTCCAGTCTGGGAATAAGTGTCAAGACGGATAAT	305
XX	12-OCT-2001; 2001US-0329690P.	Qy	427	GTGGCCACTTCTTCCCGGAGACAACCGAGATAGTCTCTGCGGATGCCAACGGAAAG	486
XX	26-OCT-2001; 2001US-0345384P.	Db	306	GTGGCCACTTCTTCCCGGAGACAACCGAGATAGTCTCTGCGGATGCCAACGGAAAG	365
XX	02-NOV-2001; 2001US-0350219P.	Qy	487	AATCTTGGGAAAGAGGCCAAACCCGAGGACACAGTCTCTAAATCTCTGTTTTTCTTGATG	546
XX	09-NOV-2001; 2001US-0344518P.	Db	366	AATCTTGGGAAAGAGGCCAAACCCGAGGACACAGTCTCTAAATCTCTGTTTTTCTTGATG	425
XX	16-NOV-2001; 2001US-0345143P.	Qy	547	CTCTCTCGGCTGTACCCAGGACGTACCGGAGACAACAGCCGAGATTCATCTCTTGGATCA	606
XX	03-DEC-2001; 2001US-0336908P.	Db	426	CTCTCTCGGCTGTACCCAGGACGTACCGGAGACAACAGCCGAGATTCATCTCTTGGATCA	485
XX	07-DEC-2001; 2001US-0340747P.	Qy	607	GTGAAGCTTGTATGTCTCAGCTCCAAATTAAGCTCCAGCGAACAAGACCCAAAGTGAGCTGG	666
XX	(INCY-) INCYTE GENOMICS INC.	Db	486	GTGAAGCTTGTATGTCTCAGCTCCAAATTAAGCTCCAGCGAACAAGACCCAAAGTGAGCTGG	545
XX	Azimzai Y, Baughn MR, Becha SD, Borowsky ML, Chawla NK;	Qy	667	ACACTTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGAGCAGCCCGG	726
XX	Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Gorvad AE;	Db	546	ACACTTCCGCTGGCAGCTGGACCGGGCAGGACACAGATAAAACCCAGGAGCAGCCCGG	605
XX	Griffin JA, Hafalia AJA, Ison CH, Kable AE, Kalafus DP;	Qy	727	GCCCAAGACAAGTCTCTCTGCGGCGCAGGATCCCAAGCTTCTCCACCTTGAGACAGGG	786
XX	Lehr-Mason PM, Lu DAM, Marquis JP, Nguyen DB, Rankumar J;	Db	606	GCCCAAGACAAGTCTCTCTGCGGCGCAGGATCCCAAGCTTCTCCACCTTGAGACAGGG	665
XX	Richardson TW, Sapperstein SK, Swarnakar A, Tang YT, Tran UK;	Qy	787	GGAGCAGGAGGAAGTCTCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTC	846
XX	Warren BA, Xu Y, Yao MG, Yue H, Yue H;	Db	666	GGAGCAGGAGGAAGTCTCTCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTC	725
XX	WPI; 2003-421159/39.	Qy	847	AAGCTGCAACAAGGACAGGAAAAAGGTGCCAGGTGACAGCCAAACAGGAAGCAAGAGGCA	906
XX	P-PSDB; ABR69628.	Db	726	AAGCTGCAACAAGGACAGGAAAAAGGTGCCAGGTGACAGCCAAACAGGAAGCAAGAGGCA	785
CC	New human proteins associated with cell growth, differentiation, and	Qy	907	GAGCATCAAGACAAGTGGATAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCCT	966
CC	death (CGPD), useful for diagnosing, treating and preventing diseases or	Db	786	GAGCATCAAGACAAGTGGATAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCCT	845
CC	conditions associated with the aberrant CGPD expression e.g. cancer,	Qy	967	GCAGGGAAGACATAGTTGACGCAAGCAAGAAAAAGAGCAAGAACTTTGGAACCTCCGAT	1026
CC	AIDS, or epilepsy.	Db	846	GCAGGGAAGACATAGTTGACGCAAGCAAGAAAAAGAGCAAGAACTTTGGAACCTCCGAT	905
CC	Claim 5; Page 330; 350pp; English.	Qy	1027	TGCTCTCTCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCTCCAGGACGCA	1086
CC	The invention relates to an isolated polypeptide associated with cell	Db	906	TGCTCTCTCTGGGACCCAGAGGACTGGAGACTGCAAGGACGATTCCTCCAGGACGCA	965
CC	growth, differentiation and death (CGPD). Also disclosed are the	Qy	1087	GCTATAGCAGAGATAATAATTCATCATGAGTTTCTTTTAAACTCTGGTTTCACTTAAC	1146
CC	polynucleotides encoding the polypeptides. The polypeptides and	Db	966	GCTATAGCAGAGATAATAATTCATCATGAGTTTCTTTTAAACTCTGGTTTCACTTAAC	1025
CC	polynucleotides are useful in diagnosing, treating and preventing	Qy	1147	AAAGCTGAAACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTCA	1206
CC	diseases or conditions associated with the decreased expression or over	Db	1026	AAAGCTGAAACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCTTCA	1085
CC	expression of CGPD. Such diseases include cell proliferative (e.g.	Qy	1207	GCTGACCTTAAGTCAGACAAGCCAACTTTATATCCAGGAGACCCAAAGGGCTGGCAAG	1266
CC	cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's	Db	1086	GCTGACCTTAAGTCAGACAAGCCAACTTTATATCCAGGAGACCCAAAGGGCTGGCAAG	1145
CC	disease, stroke), immune/inflammatory (e.g. AIDS, allergies) and	Qy	1267	AATTCGAAAGGATGCAACCCATCGGGGCACACAGTCCCGTGACAAACCCCTGAACTGCG	1326
CC	reproductive disorders, or disorders of the placenta. They are also	Db	1146	AATTCGAAAGGATGCAACCCATCGGGGCACACAGTCCCGTGACAAACCCCTGAACTGCG	1205
CC	useful in assessing the effects of exogenous compounds on the expression	Qy	1327	AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGG	1386
CC	of nucleic acid and amino acid sequences of CGPD. The CGPD or its	Db	1206	AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGG	1265
CC	fragments are useful in screening compounds for effectiveness as an	Qy			
CC	agonist or antagonist of the polypeptides, or in altering the expression	Db			
CC	of the target polynucleotide and compounds that specifically bind to or	Qy			
CC	modulate the activity of the polypeptide. Microarrays consisting	Db			
CC	polynucleotides of the invention are useful in monitoring or measuring	Qy			
CC	protein-protein interactions, drug-target interactions, and gene	Db			
CC	expression profiles. Sequences given in records ACC90578-ACC90634	Qy			
CC	represent polynucleotides encoding CGPD proteins of the invention	Db			
XX	Sequence 2020 BP; 621 A; 525 C; 529 G; 345 T; 0 U; 0 Other;	Qy			
XX	Query Match 62.9%; Score 1929.4; DB 8; Length 2020;	Db			
XX	Best Local Similarity 96.5%; Pred. No. 0;	Qy			
XX	Matches 2009; Conservative 0; Mismatches 6; Indels 66; Gaps 1;	Db			
Qy	127 GGACGAGGCTCCACCGACGCGGACTGGGACGACGCGACTGGAGACCCAGGACCCCT	186			
Db	6 GGAGCAGGCTCCACCGACGCGGACTGGGACGACGCGACTGGAGACCCAGGACCCCT	65			
Qy	187 GTGACGAGCAGCTCCCGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCA	246			
Db	56 GTGACGAGCAGCTCCCGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCA	125			
Qy	247 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGCAAGAAATGAACCA	306			

```
QY 1387 AAAAGTTCAGTTTAAAGAGAGACTCAGTCCCAAGAGTGGGAGGAGAAATGTTGGTGTGAG 1446
Db |||||||
QY 1266 AAAAGTTCAGTTTAAAGAGAGACTCAGTCCCAAGAGTGGGAGGAGAAATGTTGGTGTGAG 1325
Db |||||||
QY 1447 TCACAGTAGAGATTATTAAGTCCAAAGGAAGTAGAATCAGGCTTACAAACAGTGGACCTC 1506
Db |||||||
QY 1326 TCACAGTAGAGATTATTAAGTCCAAAGGAAGTAGAATCAGGCTTACAAACAGTGGACCTC 1385
Db |||||||
QY 1507 AACGAGAGAGATGCTGCACTGCAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAAA 1566
Db |||||||
QY 1386 AACGAAGAGAGATGCTGCTCTCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAAA 1445
Db |||||||
QY 1567 CCAAGAACCTCTCTGATGCGCTTCTCAGACAAATGTCAGTGAAGGGGATGGAGGATC 1626
Db |||||||
QY 1446 CCAAGAACCTCTCTGATGCGCTTCTCAGACAAATG----- 1481
Db |||||||
QY 1627 ACCCACTCAGAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAA 1686
Db |||||||
QY 1482 -----ACATCAGACTCCACAGAA 1499
Db |||||||
QY 1687 AAGACTATCACACCGCCAGAGCCTGAACCAACAGGAGCACACAGAGGGTAAAGAGGC 1746
Db |||||||
QY 1500 AAGACTATCACACCGCCAGAGCCTGAACCAACAGGAGCACACAGAGGGTAAAGAGGC 1559
Db |||||||
QY 1747 TCCTCGAGGACAGAGTTCAGAGCCGAGATGACACAGAGAGAGAGAGAGAGAGAGAGAA 1806
Db |||||||
QY 1560 TCCTCGAGGACAGAGTTCAGAGCCGAGATGACACAGAGAGAGAGAGAGAGAGAGAGAA 1619
Db |||||||
QY 1807 GCCAAGAACACAGCCAGTGCACAGAGCAGCGGTCGACAGCACTCACTCGAGAT 1866
Db |||||||
QY 1620 GCCAAGAACACAGCCAGTGCACAGAGCAGCGGTCGACAGCACTCACTCGAGAT 1679
Db |||||||
QY 1867 GGGGACAAAGCTCCAAAGAGACTGAGAGCGGAGAGTCCCTTGGGGGCTTCTTTAA 1926
Db |||||||
QY 1680 GGGGACAAAGCTCCAAAGAGACTGAGAGCGGAGAGTCCCTTGGGGGCTTCTTTAA 1739
Db |||||||
QY 1927 GGCCTGGGACCAAGCCGATGCTGATGCTCAAGTGCACAGAGCCAGCCAGTATCATCGGA 1986
Db |||||||
QY 1740 GGCCTGGGACCAAGCCGATGCTGATGCTCAAGTGCACAGAGCCAGCCAGTATCATCGGA 1799
Db |||||||
QY 1987 CCAGTTGGCAACCAAGTAAACAAATCAGACACGGTTCCTCCACAGAGTTCCTGCCACCA 2046
Db |||||||
QY 1800 CCAGTTGGCAACCAAGTAAACAAATCAGACACGGTTCCTCCACAGAGTTCCTGCCACCA 1859
Db |||||||
QY 2047 AGATGTGTTCTCTTACTCCATCTCTCTCCCAACACAGCTCCATGTATATATTTCTCTGA 2106
Db |||||||
QY 1860 AGATGTGTTCTCTTACTCCATCTCTCTCCCAACACAGCTCCATGTATATATTTCTCTGA 1919
Db |||||||
QY 2107 TGGCCAGCAATGAATTCGCTAGAAATTAAGCCCGAGCTGTGTATATTTAGAGTGTA 2166
Db |||||||
QY 1920 TGGCCAGCAATGAATTCGCTAGAAATTAAGCCCGAGCTGTGTATATTTAGAGTGTA 1979
Db |||||||
QY 2167 TTATTTAGCTCTCTGTTCCAGTCTTTCTGCAAAATTAACAG 2207
Db |||||||
QY 1980 TTATTTAGCTCTCTGTTCCAGTCTTTCTGCAAAATTAACAG 2020
Db |||||||
```

RESULT 9

AA512521
ID AA512521 standard; DNA; 2105 BP.

XX AC AA512521;

XX DT 04-DEC-2001 (first entry)

XX DE Gene #14 encoding novel human serine carboxypeptidase polypeptide #14.

XX KW Human; serine carboxypeptidase; immune disorder; angiogenic disorder;
KW hyperproliferative disorder; cardiovascular disorder; immunomodulatory;
KW cytostatic; cardiovascular; antiinflammatory; antiatherosclerotic; ds.

OS Homo sapiens.

XX
FN
XX
PD
XX
XX
PF
XX
XX
PR
XX
PR
XX
PA
XX
XX
PI
XX
DR
XX
XX
PT
PT
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ

Query Match 51.1%; Score 1567; DB 4; Length 2105;
Best Local Similarity 85.1%; Pred.No. 0;
Matches 1893; Conservative 0; Mismatches 190; Indels 142; Gaps 7;

```
QY 744 CTCTGCGCCAGGAGTCCACGCTTCTCCACCTGACAGAGGGGAGCAGAGAGAGC 803
Db |||||||
QY 804 TCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAGGACA 863
Db |||||||
QY 69 TCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAGGACA 128
Db |||||||
QY 864 GGAAGAGTCCAGGTGACAGCCCAAGAGAGCCAGAGGGGAGAGCATCAAGACAAGGT 923
Db |||||||
QY 129 GGAAGAGTCCAGGTGACAGCCCAAGAGAGCCAGAGGGGAGAGCATCAAGACAAGGT 188
Db |||||||
QY 924 GATAGGTTCTTGGCTTATCAGGCGAGTCCGATGATGTCCTGCGAGGAGAGCATAGT 983
Db |||||||
QY 189 GATAGGTTCTTGGCTTATCAGGCGAGTCCGATGATGTCCTGCGAGGAGAGCATAGT 248
Db |||||||
QY 984 TGACGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1043
Db |||||||
QY 249 TGACGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 308
Db |||||||
QY 1044 CCCAGAGGAGTGGAGACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1103
Db |||||||
QY 309 CCCAGAGGAGTGGAGACTGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 368
Db |||||||
QY 1104 TAATTCATCATGAGTTCTTTAAACTCTGGTTTCACTTAAACAAAGCTGAGGAGGAGGAG 1163
Db |||||||
QY 369 TAATTCATCATGAGTTCTTTAAACTCTGGTTTCACTTAAACAAAGCTGAGGAGGAGGAG 428
Db |||||||
QY 1164 GGACCCAGAGAGACCGGTCGTAAGAGTCAACCCACCTTCCAGTCACTTGAAGTCAGA 1223
Db |||||||
QY 429 GGACCCAGAGAGACCGGTCGTAAGAGTCAACCCACCTTCCAGTCACTTGAAGTCAGA 466
Db |||||||
```


Qy 1224 CAAAGCCAACTTTTACATCCACAGAGAGCCCAAGGGCTGGCAAGAAATTCCTCAAGGATGCAA 1283
Db 467 GTGATGGACAGCTGGTCAGAGACATCCGAGATCCAGGCTAGAGGCACCAAGAA---AA 523
Qy 1284 CCAATCGGGGCACACACAGTCGCTGACAAACCCCTGAAACCTGCGAAGGAAGSCACCAAGGA 1343
Db 524 AGCACTGGATAGCCCAAGGCTAGGACTCGCCTTTA----- 559
Qy 1344 GAAATCAGGACCCACTCTCTGCTCTGCGCTCTGCGGCAAACTGTTTGGAAAAAGTCAGTTAAAGA 1403
Db 560 -----GAAATTTCTTTAGGCATTAAGGGTGTCTGAAA 590
Qy 1404 GGACTCAGTCCCAAGTGGAGGAGAAATGTGGTGTGAGTCACCAAGTAGAGATTAT 1463
Db 591 GTACCCACCACTTACGTGACCTTAAG-----TCAGACAAAGCCAACTTTATCATCCAG 645
Qy 1464 AAAGTCCAAGGAAGTAGAATCAGCCTTACAAACAGTGGACCTCAACGAAGGAGATGCTGC 1523
Db 646 GAGACCCAAAGGGCTGGCA--AGAATTCCAAGGATGCAACCCATCGGGGCACACAGT 703
Qy 1524 ACCTGAACCCACAGAGGAAACTCAAAAGAGAGAAAGCAAAACCAAGAACCTCTCTGAT 1583
Db 704 CCGTGACAAACCCCTGAACCTCGAAGGAGGACCAAGGAGAAATCAGGACCCACCTCTC 763
Qy 1584 GGCCTTTCTCAGACAAATGTCAAGTGAAGGGATGGAGGGATCACCCACTCAGAGAAGAT 1643
Db 764 TGCCTCT----- 770
Qy 1644 AAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAAGACTATCACACCGCC 1703
Db 771 ---GGGCAAACTGTTTGGAAAAAGACATCAGACTCCACAGAAAAGACTATCACACCGCC 827
Qy 1704 AGAGCTGAACCAACAGAGACACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGAA 1763
Db 828 AGAGCTGAACCAACAGAGACACACAGAGGGTAAAGAGGGCTCTCGAAGGACAAAGAA 887
Qy 1764 GTACAGCGCCAGATGAACAGAGAGAGAGCAACAGCAGAGCAACCAAGAACACAGCCCA 1823
Db 888 GTACAGCGCCAGATGAACAGAGAGAGAGCAACAGCAGAGCAACCAAGAACACAGCCCA 947
Qy 1824 GTGACAGAGAGGCGCACGCTGGACACGAACTCACTGCAGAAATGGGGACAAAGCTCCAAA 1883
Db 948 GTGACAGAGAGGCGCACGCTGGACACGAACTCACTGCAGAAATGGGGACAAAGCTCCAAA 1007
Qy 1884 GAGACCTGAGAAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGCCCTGGACCAAAAGCG 1943
Db 1008 GAGACCTGAGAAAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCGCTGGACCAAAAGCG 1067
Qy 1944 GATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGACCCAGTTGGCAAAACCCAA 2003
Db 1068 GATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTGGCAAAATCCAA 1127
Qy 2004 GTAAACAAATCAGCACGCTTCCCAAGGTTCTCTGCGCAACCAAGATGTCTCTTAC 2063
Db 1128 GTAAACAAATCAGCACGCTTCCCAAGGTTCTCTGCGCAACCAAGATGTCTCTTAC 1187
Qy 2064 TCCATCTCTCTCCCAACACGCTCCATGTATATATTCTTCTGATGGCCAGCAATGAAT 2123
Db 1188 TCCATCTCTCTCCCAACACGCTCCATGTATATATTCTTCTGATGGCCAGCAATGAAT 1247
Qy 2124 TCTGCTAGAAATTAAGCCGAGCTGTTGTTATATTAGGTGTATTATTAGCTCTCTGGT 2183
Db 1248 TCTGCTAGAAATTAAGCCGAGCTGTTGTTATATTAGGTGTATTATTAGCTCTCTGGT 1307
Qy 2184 CCAGTCTTTTCTGGCAATAACAGTAAGATGTTTACAGGTCACTCAGTGTGGGTGAGA 2243
Db 1308 CCAGTCTTTTCTGGCAATAACAGTAAGATGTTTACAGGTCACTCAGTGTGGGTGAGA 1367
Qy 2244 AGAGTCGATGATCACCAGAGGAAAGGGAGGGAATAGAGGAATGTGTTCCGGTTAAGTG 2303
Db 1368 AGAGTCGATGATCACCAGAGGAAAGGGAGGGAATAGAGGAATGTGTTCCGGTTAAGTG 1427

Qy 2304 ATGAAAAATGACAGTGTGGCCGGCGTGTGGCTCTCGCTCTAATCTCAGCACTTTGGG 2363
Db 1428 ATGAAAAATGACAGTGTGGCCGGCGTGTGGCTCTCGCTCTAATCTCAGCACTTTGGG 1487
Qy 2364 AGCCGAGGAGGAGGTGATCACTGAGGTCAAGGTCAAGACTAGCCTGGCCCAACATCAT 2423
Db 1488 AGCCGAGGAGGAGGTGATCACTGAGGTCAAGGTCAAGACTAGCCTGGCCCAACATCAT 1547
Qy 2424 GAAACCCCGTCTCTACTTAAAAATACAAAAATTAAGCCAGGCATGGTGGCACACACCTGTAG 2483
Db 1548 GAAACCCCGTCTCTACTTAAAAATACAAAAATTAAGCCAGGCATGGTGGCACACACCTGTAG 1607
Qy 2484 TCCAGCTACTCGGGAGCCCAAGCCAGACGAGAACCGCTTGTACCCAGAGGTGGAGTTGC 2543
Db 1608 TCCAGCTACTCGGGAGCCCAAGCCAGACGAGAACCGCTTGTACCCAGAGGTGGAGTTGC 1667
Qy 2544 AGTGAGCCGAAGTTGCAACCAATTGCACTTCAACCCCTGGGCGACAGAGCAAGATTCTATC-AA 2602
Db 1668 AGTGAGCCGAAGTTGCAACCAATTGCACTTCAACCCCTGGGCGACAGAGCAAGATTCTATCAAA 1727
Qy 2603 AAAAAAAGGCGAGTGGCAAGTATAGAGAGAAATGCTGCTAGAAGGAATTAAGCG 2662
Db 1728 AAAAAAAGGCGAGTGGCAAGTATAGAGAGAAATGCTGCTAGAAGGAATTAAGCG 1787
Qy 2663 TTGTAGTAAACGGGTGCTCTCTCTCTAAGCTTGAAGAGGGAGACGAAAAATCCATTTGTT 2722
Db 1788 TTGTAGTAAACGGGTGCTCTCTCTCTCTAAGCTTGAAGAGGGAGACGAAAAATCCATTTGTT 1847
Qy 2723 TAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTTGGGTGGTGGCCAAATTTCTAG 2782
Db 1848 TAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTTGGGTGGTGGCCAAATTTCTAG 1907
Qy 2783 AACCGAATGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAATAGGTTCC 2842
Db 1908 AACCGAATGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAATAGGTTCC 1967
Qy 2843 TTGTAACTTTATGATGAGAGGAAAAAGATTGACTGGGGAGGCTTAAAAATGATTGGGAA 2902
Db 1968 TTGTAACTTTATGATGAGAGGAAAAAGATTGACTGGGGAGGCTTAAAAATGATTGGGAA 2027
Qy 2903 AACAAATGCTTTGAGGCTCAGTGACAAACCGCAAGATTACAACTTAAAAAATAAAAAA 2962
Db 2028 AACAAATGCTTTGAGGCTCAGTGACAAACCGCAAGATTACAACTTAAAAAATAAAAAA 2087
Qy 2963 AAAAA 2967
Db 2088 AAAAA 2092

RESULT 10
ABQ77258
ID ABQ77258 standard; cdna; 2105 BP.
XX
AC ABQ77258;
XX
DT 17-APR-2003 (first entry)
XX
DE Human tenascin-W homologue HSDET79 cdna.
XX
KW Human; antiallergic; antiasthmatic; antiinflammatory; antibacterial;
KW virucide; fungicide; antidiabetic; anti-HIV; cytostatic; vulnerary;
KW antianemic; antiarteriosclerotic; immunosuppressive; neotropic; AIDS;
KW neuroprotective; antirheumatic; vasotropic; antiarthritic; antithyroid;
KW dermatological; antiparkinsonian; cancer; hyperproliferative disorder;
KW autoimmune disorder; systemic lupus erythematosus; multiple sclerosis;
KW haematopoietic disorder; haematologic disorder; eczema; gene therapy; ss;
KW inflammatory bowel disease; Crohn's disease; neurodegenerative disorder;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder; HIV;
KW infectious disease; wound healing; epithelial cell proliferation; gene.
XX Homo sapiens.
OS
XX
FN US2002151009-A1.

PD 17-OCT-2002.
XX 28-AUG-2001; 2001US-00939825.
XX 24-FEB-2000; 2000US-0184664P.
PR 16-MAR-2000; 2000US-0189874P.
PR 22-FEB-2001; 2001WO-US005498.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ni J, Shi Y, Ebner R, Ruben SM;
PI WPI: 2003-198289/19.
XX P-PSDB; AB674534.
DR
DR
XX
PT New human polypeptides and nucleic acid molecules, useful for detecting,
PT preventing, diagnosing, prognosticating, treating or ameliorating medical
PT conditions such as cancer, AIDS, Alzheimer's disease or Parkinson's
PT disease.
XX
PS Claim 1; Page 137-138; 152pp; English.
XX
CC This invention describes novel human polypeptide and the polynucleotides
CC that encode them which have anti-allergic, antiasthmatic, antibacterial,
CC anti-inflammatory, virucide, fungicide, antidiabetic, anti-HIV, nootropic,
CC cytostatic, vulnerary, antianemic, antiarteriosclerotic, antiarthritic,
CC immunosuppressive; neuroprotective, antirheumatic, vasotropic,
CC dermatological, antithyroid and antiparkinsonian activity. The
CC polypeptides and nucleic acid molecules are useful for detecting,
CC preventing, diagnosing, prognosticating, treating or ameliorating medical
CC conditions such as cancer or other hyperproliferative disorders,
CC autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus
CC erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic
CC anaemia), haematopoietic or haematologic disorders (e.g. anaemia,
CC thrombocytopenia), allergic reactions including asthma or eczema,
CC inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory
CC bowel disease or Crohn's disease), neurodegenerative disorders (e.g.
CC Alzheimer's disease or Parkinson's disease), cardiovascular disorders
CC (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial,
CC fungal or viral infections including HIV/AIDS), or wound healing and
CC disorders of epithelial cell proliferation. The nucleic acids are also
CC useful for gene therapy, chromosome identification, radiation hybrid
CC mapping or long-range restriction mapping, as molecular weight markers,
CC or as hybridisation or diagnostic probes. The polypeptides and antibodies
CC are useful for providing immunological probes for differential
CC identification of the tissues immunohistochemistry assays. The methods
CC are also useful for inhibiting or enhancing the production and function
CC of the polypeptide, or identifying a binding partner for the polypeptide.
CC This sequence encodes a polypeptide described in the disclosure of the
CC invention
XX
SQ Sequence 2105 BP; 678 A; 479 C; 553 G; 395 T; 0 U; 0 Other;
Query Match 51.1%; Score 1567; DB 10; Length 2105;
Best Local Similarity 85.1%; Pred. No. 0;
Matches 1893; Conservative 0; Mismatches 190; Indels 142; Gaps 7;
744 CTCCTCCCGCAGGGATCCACGCTTCTCCACCTGAGACAGGGGAGGAGGAGGAGC 803
9 CTCCTCCCGCAGGGATCCACGCTTCTCCACCTGAGACAGGGGAGGAGGAGGAGC 68
804 TCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGCAAGGGACA 863
69 TCCCTCCAGGCCCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGCAAGGGACA 128
864 GGAAAGGTGCCAGGTGACAGCCCAAGAGCCCAAGAGGGGAGGAGGAGGAGGAGGAGT 923
129 GGAAAGGTGCCAGGTGACAGCCCAAGAGGCCCAAGAGGGGAGGAGGAGGAGGAGGAGT 188
924 GGATCAGGTTCTTGGCTTATCAGGGCAGTCCGATGATCTCCCTGCGAGGAGGAGGAGT 983
189 GGATCAGGTTCTTGGCTTATCAGGGCAGTCCGATGATCTCCCTGCGAGGAGGAGGAGT 248

QY 984 TGACGGCAAGGAAAGGAGGAGCAAGAACTTGGAACTGCGGATTCCTCTGTCCCTGGGGA 1043
DB 249 TGACGGCAAGGAAAGGAGGAGCAAGAACTTGGAACTGCGGATTCCTCTGTCCCTGGGGA 308
QY 1044 CCCAGAAGGACTGGAGACTGCAAAAGGAGGAGTCCAGGAGCAGCAGCTATAGCAGAGAAATAA 1103
DB 309 CCCAGAAGGACTGGAGACTGCAAAAGGAGGAGTCCAGGAGCAGCAGCTATAGCAGAGAAATAA 368
QY 1104 TAATTCCATCATGAGTTTCTTTAAACACTCTGTTTCACTAAACAGCTGAAACAAAAA 1163
DB 369 TAATTCCATCATGAGTTTCTTTAAACACTCTGTTTCACTAAACAGCTGAAACAAAAA 428
QY 1164 GGACCCAGAAGACAGGGTGTCTGAAAAGTCAACCCACACTTCAGCTGACCTTAACTCAGA 1223
DB 429 GGACCCAGAAGACAGGGTGTCTGAAAAGTCAACCCACACTTCAGCTGACCTTAACTCAGA 466
QY 1224 CAAAGCCAACTTTACATCTCCAGGAGAGCCCAAGGGCTGCGAAGATTTCCAAAGGATGCAA 1283
DB 467 GTGATGGACAGCTGGTTCAGAAAGCATCCGAGATCCAGGCTAGAGGCAACCAAGAA---AA 523
QY 1284 CCCATCGGGGCACACACAGTCCGTGACAAACCCCTGAACCTCGGAGGAGGAGCACCAGGA 1343
DB 524 AGCACTCGATAGCCCGAGGCTAGGACTCGCCTTTA----- 559
QY 1344 GAAATCAGGACCCACCTCTCTGCGCTCTGGGCAAACTGTTTGGAAAAAGTCAAGTTAAAGA 1403
DB 560 -----GAAATTTCTTTAGGCAATAAGGGTGTCTGAAAA 590
QY 1404 GGACTCAGTCCCCACAGTGGGGAGGAGAAATGTTGTTGTTGAGTCCACAGTAGAGATTAT 1463
DB 591 GTCACCCACCACTTCAGCTGACCTTAAG-----TCAGACAAAGGCCAATTTACATCCAG 645
QY 1464 AAGTCCAAAGGAAGTAGAATCAGCTTCAAAACAGTGCACCTCAACGAGGAGAGATGCTGC 1523
DB 646 GAGACCCAAAGGGCTGGCA--AGATTTCCAAAGGATGCAACCCATCGGGGCACACACAGT 703
QY 1524 ACCTGAACCCACAGAGGCAAACTCAAAAGAGAGAAAGCAAAACCAAGAACTCTCTGAT 1583
DB 704 CCGTGACAAACCCCTGAACCTCGAAGGAGGAGCAACCAAGGAGAAATTCAGGACCCACCTCTC 763
QY 1584 GGCGTTTCTCAGACNAATGTTCAGTGAAGGGGATGAGGGATCACCCACTCAGAGAAAT 1643
DB 764 TGCCTCT----- 770
QY 1644 AAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATCACACCGCC 1703
DB 771 ---GGGCAAACTGTTTGGAAAAAGACATCAGACTCCACAGAAAGACTATCACACCGCC 827
QY 1704 AGAGCTTGAAACCAACAGGAGCAACAGAGGGTAAAGAGGGCTCTCGAAGGACAAGAA 1763
DB 828 AGAGCTTGAAACCAACAGGAGCAACAGAGGGTAAAGAGGGCTCTCGAAGGACAAGAA 887
QY 1764 GTGACAGCCGAGATGACAGAGGAGAGAGCAACAGGAGGAGCCAAAGAACCCAGCCCA 1823
DB 888 GTCAGAGCCGAGATGAACAGAGAGAGAGCAACAGGAGGAGCCAAAGAACCCAGCCCA 947
QY 1824 GTGACAGAGCAGGCGCCAGGTGGACAGAACTCACTGAGAAATGGGACAAAGCTCCAAAA 1883
DB 948 GTGACAGAGCAGGCGCCAGGTGGACAGAACTCACTGAGAAATGGGACAAAGCTCCAAAA 1007
QY 1884 GAGACTTGAGAAAGCGGAGCAGTCCCTTGGGGGCTTTTAAAGGCTCTGGGACCAAGCG 1943
DB 1008 GAGACTTGAGAAAGCGGAGCAGTCCCTTGGGGGCTTTTAAAGGCTCTGGGACCAAGCG 1067
QY 1944 GATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGACAGTGTGGCAACCCAA 2003
DB 1068 GATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGACAGTGTGGCAACCCAA 1127
QY 2004 GTAAACAAATCAGCACCGTTCCACAGGTTCTCTGCGACCAAGATGTGTTCTCTTAC 2063
DB 1128 GTAAACAAATCAGCACCGTTCCACAGGTTCTCTGCGACCAAGATGTGTTCTCTTAC 1187

```
Qy 2064 TCCATCTCCTCCCAACACGCTCCATGTATATATTTCTTGTATGGCCAGCAAAATGAAT 2123
Db 1188 TCCATCTCCTCCCAACACGCTCCATGTATATATTTCTTGTATGGCCAGCAAAATGAAT 1247
Qy 2124 TCTCCCTAGAAATTAAGCCCGAGCTGTTGTATATTTAGGTTGTATATTACGTTCTCTGGT 2183
Db 1248 TCTCCCTAGAAATTAAGCCCGAGCTGTTGTATATTTAGGTTGTATATTACGTTCTCTGGT 1307
Qy 2184 CCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTACAGGTCACCTAGTTGGGTGAG 2243
Db 1308 CCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTACAGGTCACCTAGTTGGGTGAG 1367
Qy 2244 AGAGTCGATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2303
Db 1368 AGAGTCGATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1427
Qy 2304 ATGAAATGGCAGTGGTGGCGGGGGGGTGGTGGCTCTGGCCTGTAATCTCAGCATTGGG 2363
Db 1428 ATGAAATGGCAGTGGTGGCGGGGGGGTGGTGGCTCTGGCCTGTAATCTCAGCATTGGG 1487
Qy 2364 AGGCGGAGGAGGAGGATCAGCTCAGGTCAGGAGTTCAAGACTAGCCTGGCCACATCAT 2423
Db 1488 AGGCGGAGGAGGAGGATCAGCTCAGGTCAGGAGTTCAAGACTAGCCTGGCCACATCAT 1547
Qy 2424 GAAACCCCGTCTCTACTAAATAACAAATAATAGCCAGGCTAGTGGGACACACCTGTAG 2483
Db 1548 GAAACCCCGTCTCTACTAAATAACAAATAATAGCCAGGCTAGTGGGACACACCTGTAG 1607
Qy 2484 TCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGTAACCCAGGAGTGGAGGTTGC 2543
Db 1608 TCCAGCTACTCGGGAGGCCAACGACGAGAACCGCTTGTAACCCAGGAGTGGAGGTTGC 1667
Qy 2544 AGTGAGCCGAGTTGCACATTTGCATCCACCTGGGCGACAGAGCAAGATTTCTATC-AA 2602
Db 1668 AGTGAGCCGAGTTGCACATTTGCATCCACCTGGGCGACAGAGCAAGATTTCTATCAA 1727
Qy 2603 AAAAAAAGGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGCTAGAGGAATTAAGCG 2662
Db 1728 AAAAAAAGGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGCTAGAGGAATTAAGCG 1787
Qy 2663 TTGTAGTAAACGGTGTCTCATCTCTAAGCTTGAAGAGGAGAGCAAAATCCATTTGTT 2722
Db 1788 TTGTAGTAAACGGTGTCTCATCTCTAAGCTTGAAGAGGAGAGCAAAATCCATTTGTT 1847
Qy 2723 TAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTTGGGTGTTGCCAAATTTCTAG 2782
Db 1848 TAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTTGGGTGTTGCCAAATTTCTAG 1907
Qy 2783 AACGGAATGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAAATAGGGTCC 2842
Db 1908 AACGGAATGTGGGGTATAGAAAAAGGAATGAATAAGCGTTGTTTTTCAAATAGGGTCC 1967
Qy 2843 TTGTAAGTTATGTAGAGGAGGAAAGATTGACTGGGGAGGCTTAAATGATTTGGGAA 2902
Db 1968 TTGTAAGTTATGTAGAGGAGGAAAGATTGACTGGGGAGGCTTAAATGATTTGGGAA 2027
Qy 2903 AACAAATCTTTTGGGGTCAAGTGCACACGCAAGATTACAACCTTAAAAAATAAAAAA 2962
Db 2028 AACAAATCTTTTGGGGTCAAGTGCACACGCAAGATTACAACCTTAAAAAATAAAAAA 2087
Qy 2963 AAAAA 2967
Db 2088 AAAAA 2092
```

RESULT 11

ABAl7779

ID ABAl7779 standard; DNA; 9899 BP.

XX

AC ABAl7779;

XX

DT 23-JAN-2002 (first entry)

XX

Human nervous system related polynucleotide SEQ ID NO 10110.
Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
antiparkinsonian; antiskinking; antianaemic; antiarthritic; cancer;
antirheumatic; hepatotropic; cerebroprotective; antinflammatory;
antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
antiparasitic; cardiac; immune disorder; cardiovascular disorder;
neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.

Homo sapiens.

WO200159063-A2.

16-AUG-2001.

17-JAN-2001; 2001WO-US001334.

31-JAN-2000; 2000US-0179065P.

04-FEB-2000; 2000US-0180628P.

24-FEB-2000; 2000US-0184664P.

02-MAR-2000; 2000US-0186350P.

16-MAR-2000; 2000US-0189874P.

17-MAR-2000; 2000US-0190076P.

18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.

07-JUN-2000; 2000US-0209467P.

28-JUN-2000; 2000US-0214886P.

30-JUN-2000; 2000US-0215135P.

07-JUL-2000; 2000US-0216647P.

07-JUL-2000; 2000US-0216880P.

11-JUL-2000; 2000US-0217487P.

14-JUL-2000; 2000US-0217496P.

14-JUL-2000; 2000US-0218290P.

26-JUL-2000; 2000US-0220963P.

26-JUL-2000; 2000US-0220964P.

14-AUG-2000; 2000US-0224518P.

14-AUG-2000; 2000US-0224519P.

14-AUG-2000; 2000US-0225213P.

14-AUG-2000; 2000US-0225214P.

14-AUG-2000; 2000US-0225266P.

14-AUG-2000; 2000US-0225267P.

14-AUG-2000; 2000US-0225268P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

14-AUG-2000; 2000US-0225270P.

PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241788P.
PR 20-OCT-2000; 2000US-0241789P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 20-OCT-2000; 2000US-0242221P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.

PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.
XX Disclosure; SEQ ID NO 10110; 1701pp + Sequence Listing; English.
XX The invention relates to novel genes (AB11004-ABA21534) and proteins
CC (AB114678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 9889 BP; 2944 A; 2085 C; 2204 G; 2656 T; 0 U; 0 Other;
Best Local Similarity 33.2%; Score 1016.4; DB 5; Length 9889;
Query Match 99.1%; Pred. No. 9.4e-246;
Matches 1033; Conservative 0; Mismatches 6; Indels 3; Gaps 1;
Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACCAACAGACCCAGTATCCATCGGACC 1988
Db |||||
Qy 5062 CCAGGGACCAAGCGGATGTTGGATGCTCAAGTGCACCAACAGACCCAGTATCCATCGGACC 5121
Db |||||
Qy 1989 AGTTGGCAACCCCAAGTAACAAATCAGACGGTTCACCAAGTTCCTCGCCACCAAG 2048
Db |||||
Qy 5122 AGTTGGCAATCCCAAGTAACAAATCAGACGGTTCACCAAGTTCCTCGCCACCAAG 5181
Db |||||
Qy 2049 ATGTGTTCTCCTTACTCCTCTCCCAACAGCTCCATGTATATATTTCTCTGATG 2108
Db |||||
Qy 5182 ATGTGTTCTCCTTACTCCTCTCCCAACAGCTCCATGTATATATTTCTCTGATG 5241
Db |||||
Qy 2109 GCCAGCAAAATGAAATTTCTGCTAGAAATTAAGCCGAGCTCTGTATATTCAGGTGATT 2168
Db |||||
Qy 5242 GCCAGCAAAATGAAATTTCTGCTAGAAATTAAGCCGAGCTCTGTATATTCAGGTGATT 5301
Db |||||
Qy 2169 ATTTACGTCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTAGCAGGTCA 2228
Db |||||
Qy 5302 ATTTACGTCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTAGCAGGTCA 5361
Db |||||
Qy 2229 CCTAGTTGGGTCAGAAGTGCATGATCACCACAGGAAGGGAGGAATAGAGGAATG 2288
Db |||||
Qy 5362 CCTAGTTGGGTCAGAAGTGCATGATCACCACAGGAAGGGAGGAATAGAGGAATG 5421
Db |||||
Qy 2289 TGTTCCGGGTTAAGTCATGAAATGGCAGTGTGGCGGGCGGTGGTGGCTCTCGCCTGTAA 2348
Db |||||
Qy 5422 TGTTCCGGGTTAAGTCATGAAATGGCAGTGTGGCGGGCGGTGGTGGCTCTCGCCTGTAA 5481
Db |||||

```
Qy 2349 TCTCAGCACTTTGGAGCCGAGGCGAGTGATGATCACTGAGGTTCAGAGTCAAGACTAG 2408
Db 5482 TCTCAGCACTTTGGAGCCGAGGCGAGTGATGATCACTGAGGTTCAGAGTCAAGACTAG 5541
Qy 2409 CCTGCCACATCATGAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGCATGGT 2468
Db 5542 CCTGCCACATCATGAACCCCGTCTCTACTAAAAATACAAAAATTAGCCAGCATGGT 5601
Qy 2469 GGACACACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCA 2528
Db 5602 GGACACACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCA 5661
Qy 2529 GGAGGTGAGGTTGCAGTGAGCCGAAGTTGCACCAATTGCACTCCACCCCTGGCGACAGAG 2588
Db 5662 GGAGGTGAGGTTGCAGTGAGCCGAAGTTGCACCAATTGCACTCCACCCCTGGCGACAGAG 5721
Qy 2589 CAAGATTCATC-----AAAAAAAAAGGCGAGTGCGCAAGTAAATTAGAGAGAAATGCTG 2645
Db 5722 CAAGATTCATCAAAAAAAAAAAAAAAAAAGGCGAGTGCGCAAGTAAATTAGAGAGAAATGCTG 5781
Qy 2646 CTAGAAGGAATTAAGCGTTGTAGTAAACCGCTGCTCATCTCTAAGCTTGAAGAAGGGAG 2705
Db 5782 CTAGAAGGAATTAAGCGTTGTAGTAAATGCGTCTTATCTCTAAGCTTGAAGAAGGGAG 5841
Qy 2706 ACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGTTGGGT 2765
Db 5842 ACGAAATCCATTTGTTTAAATTCACATCTCAAGGAGGGAGAACCCGGGCTGTGTTGGGT 5901
Qy 2766 GGTGGCAATTTCTAGAACGGAATGTGTTGGGTATAGAAAAAGGAATGAATAGCGTTG 2825
Db 5902 GGTGGCAATTTCTAGAACGGAATGTGTTGGGTATAGAAAAAGGAATGAATAGCGTTG 5961
Qy 2826 TTTTTCAAATAGGTCCTTTGAAGTTATGATGAGAGGAGAAAGATTGACTGGGGAGGGC 2885
Db 5962 TTTTTCAAATAGGTCCTTTGAAGTTATGATGAGAGGAGAAAGATTGACTGGGGAGGGC 6021
Qy 2886 TTAATAATGATTGGGAAAAACAATTGCTTTTGGAGGCTCAGTGACAAACGCAAGATTACAA 2945
Db 6022 TTAATAATGATTGGGAAAAACAATTGCTTTTGGAGGCTCAGTGACAAACGCAAGATTACAA 6081
Qy 2946 CTTAAAAAATAAAAAAAAAAAAAA 2967
Db 6082 CTTAAAAAATAAAAAAAAAATAA 6103
```

RESULT 12

ABZ36099/C

ID ABZ36099 standard; cDNA; 1504 BP.

XX AC

XX ABZ36099;

XX DT 10-FEB-2003 (first entry)

XX DE Human secretory polynucleotide SPTM SEQ ID NO 263.

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX KW

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

XX OS

PR 29-MAR-2001; 2001US-0280068P.

PR 16-MAY-2001; 2001US-0291280P.

PR 17-MAY-2001; 2001US-0291829P.

PR 17-MAY-2001; 2001US-0291849P.

PR 19-JUN-2001; 2001US-0299428P.

PR 20-JUN-2001; 2001US-0299776P.

PR 20-JUN-2001; 2001US-030001P.

XX (INCY-) INCYTE GENOMICS INC.

XX Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;

XX Dufour GE, Hillman JL, Yu JY, Tuason O, Yap PE, Amshy SR;

XX Dugherly SC, Dam TC, Liu TF, Nguyen DA, Kleefeld Y, Gerstein EH;

XX Peralta CH, David MH, Lewis SA, Chen AJ, Panzer SR, Harris B;

XX Flores V, Marwaha R, Lo A, Lan RY, Urashka ME;

XX WPI; 2003-075543/07.

XX P-PSDB; ABP75656.

XX New human secretory proteins and polynucleotides, useful for diagnosing,

XX treating or preventing autoimmune/inflammatory disorders (e.g. AIDS),

XX neurological disorders (e.g. Alzheimer's), or cell proliferations or

XX cancers.

XX Claim 1; SEQ ID NO 263; 458pp + Sequence Listing; English.

XX The invention relates to a secretory polynucleotide (designated sptm)

XX comprising any of 567 polynucleotide sequences (ABZ35837-ABZ36403), a

XX naturally occurring polynucleotide sequence at least 90 % identical to

XX the polynucleotide sequence, a polynucleotide complementary to them or an

XX RNA equivalent of them. The polypeptide or polynucleotide are useful for

XX treating, preventing or diagnosing a disease or condition associated with

XX the expression of functional SPTM. These are particularly useful for

XX diagnosing, treating or preventing autoimmune/inflammatory disorders

XX (e.g. acquired immunodeficiency syndrome, anaemia, asthma or Crohn's

XX disease), neurological disorders (e.g. epilepsy, Huntington's disease,

XX dementia, stroke, Alzheimer's disease, Creutzfeldt-Jakob disease,

XX multiple sclerosis, cerebral palsy, Parkinson's disease, anxiety,

XX schizophrenia or amnesia), or cell proliferative disorders (e.g.

XX psoriasis, polycythemia vera, or cancers including adenocarcinoma,

XX leukaemia, lymphoma, melanoma, myeloma, sarcoma or cancers of the brain,

XX breast, cervix or prostate). Note: The sequence data for this patent did

XX not form part of the printed specification, but was obtained in electronic

XX format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ

SQ Sequence 1504 BP; 326 A; 394 C; 399 G; 385 T; 0 U; 0 Other;

Query Match 19.0%; Score 584; DB 8; Length 1504;

Best Local Similarity 93.8%; Pred. No. 7,3e-137;

Matches 630; Conservative 0; Mismatches 40; Indels 2; Gaps 2;

Qy 2233 GTTGGGTCAAGAGTTCGATGATCAACCAAGCAGGAGGAGGAAATAGAGGAATGTGTT 2292

Db 670 GTGGGACCGACTGTATCAACATCACCAGCAGGAGGAGGAAATAGAGGAATGTGTT 611

Qy 2293 CGGTTAAGTATGAAATGCGAGTGTGGCGGCGTGTGGCTCTCGCTCTGAATCTC 2352

Db 610 CGGTTAAGTATGAAATGCGAGTGTGGCGGCGTGTGGCTCTCGCTCTGAATCTC 551

Qy 2353 AGCACTTTGGAGCCCGAGGCGAGTGTGATCACCTGAGGTACGAGTTCAAGACTAGCTG 2412

Db 550 AGCGTTTGGAGCCCGAGGCGGTTGGATCACCTGAGGTTCGGAGTTCAAGACTAGCTG 491

Qy 2413 GCCAACATCATGAACCCCGTCTCTACTTAAATAATACAAATTAAGCAGGATCGTGCA 2472

Db 490 GCCAACATCATGAACCCCGTCTCTACTTAAATAATACAAATTAAGCAGGATCGTGCA 431

Qy 2473 CACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAG 2532

Db 430 CACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGTACCCAGGAG 371

Qy 2533 GTGAGGTTGAGTGAGCCGGAAGTTGACCATTCACCTCCACCTGGCGGACGAGCAAG 2592

Db 370 GTGGAGGTTGCGTGAGCGGAAGTTGACCATTTGACCTCCACCTGGCGGACAGCAAG 311
QY ACH89479/c
XX ACH89479 standard; DNA; 644 BP.
AC ACH89479;
XX
XX
DT 29-JUL-2004 (first entry)
XX Human genome derived single exon probe #22674.
XX Human; probe; ss; gene expression; single exon probe; microarray;
KW alternative splicing event; genomic alteration.
XX Homo sapiens.
XX US2003194704-A1.
XX 16-OCT-2003.
XX 03-APR-2002; 2002US-00029386.
XX 03-APR-2002; 2002US-00029386.
XX (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX Penn SG, Rank DR, Hanzel DK;
XX WPI; 2004-119264/12.
XX
XX New human genome-derived single exon nucleic acid probes useful for human
PT gene expression analysis, for identifying or characterizing alternative
PT splicing events, for assessing genomic alterations or as tools for
PT surveying tissues.
XX
XX Claim 1; SEQ ID NO 22674; 80pp; English.
XX
XX The invention relates to a nucleic acid probe for measuring human gene
CC expression, comprising any of the 27,400 fully defined nucleotide
CC sequences in the specification, or their complements or fragments, and
CC encoding at least 8 amino acids of any of the 6888 amino acid sequences
CC fully defined in the specification. The probe is a single exon probe that
CC hybridises under high stringency conditions to a nucleic acid molecule
CC expressed in human cells or tissues. Also included are a spatially-
CC addressable set of single exon nucleic acid probes for measuring human

CC gene expression (comprising a plurality of single exon nucleic acid
CC probes cited above, where each of the plurality of probes is separately
CC and addressably isolatable or amplifiable from the plurality), a single
CC exon microarray for measuring human gene expression, a method of
CC measuring human gene expression, a vector comprising the single exon
CC probe cited above, an ORF-encoded peptide comprising at least 8
CC contiguous amino acids of any of the above-mentioned amino acid
CC sequences (optionally with conservative amino acid substitutions), an
CC isolated antibody that binds specifically to a peptide cited above,
CC methods of selling and/or licensing single exon probes or microarrays to
CC a customer desiring to measure gene expression, a method of providing
CC human gene expression data by subscription, and a computer-readable
CC storage medium which contains a database having a plurality of records
CC (each record including data on the expression of a single exon probe
CC cited above). The probe, methods and apparatus are useful in gene
CC expression analysis. The probes may be used as tools for surveying
CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 644 BP; 118 A; 186 C; 168 G; 172 T; 0 U; 0 Other;

Query Match 19.0%; Score 582.6; DB 12; Length 644;
Best Local Similarity 99.3%; Pred. No. 1.2e-136;
Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 395 TCGACTTGGGAATAGTGTCAAGACGGATATATGTGGCCACTTCTTCCCCCGAGCAACGG 454
DB 644 TCGACTTGGGAATAGTGTCAAGACGGATATATGTGGCCACTTCTTCCCCCGAGCAACGG 585
QY 455 AGATAAGTGTCTGTGGGATGCCAACGGAAGAATCTTGGGAAAGAGGCCAAACCCGAGG 514
DB 584 AGATAAGTGTCTGTGGGATGCCAACGGAAGAATCTTGGGAAAGAGGCCAAACCCGAGG 525
QY 515 CACCAGCTGTAAATCTCGTTTCTTGTAGTCTCTCTCGGCTGTACAGGACGTACCG 574
DB 524 CACCAGCTGTAAATCTCGTTTCTTGTAGTCTCTCTCGGCTGTACAGGACGTACCG 465
QY 575 GAGACCAAGCGCAGATTTCATCCCTTGATTCAGTGAAGCTTGTATGTCAGTCCCAATAAG 634
DB 464 GAGACCAAGCGCAGATTTCATCCCTTGATTCAGTGAAGCTTGTATGTCAGTCCCAATAAG 405
QY 635 CTCCAGCGAACAAGACCCCAAGTGAGAGCTGGACACTTTCGGTGGCAGCTGGACCGGGGC 694
DB 404 CTCCAGCGAACAAGACCCCAAGTGAGAGCTGGACACTTTCGGTGGCAGCTGGACCGGGGC 345
QY 695 AGGACACAGATATAAACCCCGAGGCAACCGCCCGCCCAAGACAAGGTCCTCTCTGCGCGCA 754
DB 344 AGGACACAGATATAAACCCCGAGGCAACCGCCCGCCCAAGACAAGGTCCTCTCTGCGCGCA 285
QY 755 GGGATCCACGCTTCTCCCACTGAGACACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 814
DB 284 GGGATCCACGCTTCTCCCACTGAGACACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 225
QY 815 CCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAAAGGACAGGAAAGGTGC 874
DB 224 CCAAGGACTCCAGCTTTTGTGACAAATTTCTTCAAGCTGGACAAAGGACAGGAAAGGTGC 165
QY 875 CAGGTGACAGCCAAACAGGAAGCCAAAGAGGGCAGAGCATCAAGACAAGGTGATGAGGTTTC 934
DB 164 CAGGTGACAGCCAAACAGGAAGCCAAAGAGGGCAGAGCATCAAGACAAGGTGATGAGGTTTC 105
QY 935 CTGCTTTATCAGGCGAGTCCGATGATGTCCTCGCAGGAGGAGCATAGT 983
DB 104 CTGCTTTATCAGGCGAGTCCGATGATGTCCTCGCAGGAGGAGGTAAGTGT 56

Query Match	Best Local Similarity	Score	DB 4	Length	488;
Matches	47;	Conservative	0;	Mismatches	1;
Indels	0;	Gaps	0;		
Qy	2142	CCGAGCTGTTGTATATTGAGGTGATTATTTACGTCCTCTGTGTGTCAGTCTTTTCTGGCAAA	2201		
Db	15	CCGAGCTGTTGTATATTGAGGTGATTATTTACGTCCTCTGTGTGTCAGTCTTTTCTGGCAAA	74		
Qy	2202	TACAGTAAAGATGGTTTAGCAGTCCACCTAGTTGGGTCCAGAGAGTCGATGATCACC	2261		
Db	75	TACAGTAAAGATGGTTTAGCAGTCCACCTAGTTGGGTCCAGAGAGTCGATGATCACC	134		
Qy	2262	GCAGGAAGGAGGGAATAGAGGAATGTGTTCCGGTAAAGTGATGAAAAATGCCAGTGGTG	2321		
Db	135	GCAGGAAGGAGGGAATAGAGGAATGTGTTCCGGTAAAGTGATGAAAAATGCCAGTGGTG	194		
Qy	2322	GCCGGGCGTGTGGCTCTCGCTGTAACTCAGCACTTTTGGGAGGCCGAGCAGGTGGAT	2381		
Db	195	GCCGGGCGTGTGGCTCTCGCTGTAACTCAGCACTTTTGGGAGGCCGAGCAGGTGGAT	254		
Qy	2382	CACCTGAGGTCAGAGGTTCAAGATAGCTCGSCCAAATCATGAAACCCCGCTCTCTACTA	2441		
Db	255	CACCTGAGGTCAGAGGTTCAAGATAGCTCGSCCAAATCATGAAACCCCGCTCTCTACTA	314		
Qy	2442	AAAAATACAAAAATTAGCCAGCATGGTGGCAACACCTGTGTAGTCCAGCTACTCGGGAGC	2501		

CC tissues to detect the presence of expressed messages that contain their
CC specific exon, or in constructing genome-derived single exon microarrays.
CC In addition, the probes are used in identifying and characterising
CC alternative splicing events, in detecting and characterising gross
CC alterations in the genomic locus that includes their exon, in assessing
CC smaller genomic alterations, in priming the synthesis of nucleic acids,
CC or in expressing the ORF-encoded peptide. The present sequence is a human
CC single exon probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030194704
XX
SQ Sequence 576 BP; 151 A; 119 C; 136 G; 170 T; 0 U; 0 Other;

Query Match 10.9%; Score 332.8; DB 12; Length 576;
Best Local Similarity 99.4%; Pred. No. 1.4e-73;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 1988
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
336 CCAGGGACCAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACC 277
Qy 1989 AGTTGGCAAAACCAAGTAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCCACCAAG 2048
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
276 AGTTGGCAAAATCCAAGTAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCCACCAAG 217
Qy 2049 ATGTGTTCTCTTACTCCATCTCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 2108
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
216 ATGTGTTCTCTTACTCCATCTCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 157
Qy 2109 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGATT 2168
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
156 GCCAGCAATGAAATTTCTGCCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGATT 97
Qy 2169 ATTTACGTCCTCGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 2228
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
96 ATTTACGTCCTCGTCCAGTCTTTTCTGGCAAAATAACAGTAAAGATGGTTTAGCAGGTCA 37
Qy 2229 CCTAGTTGGGTCAGAAGAGTCGATGATCACCAGCA 2264
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
36 CCTAGTTGGGTCAGAAGAGTCGATGATCACCAGCA 1

Search completed: July 3, 2005, 10:07:13
Job time : 1575.51 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 483.521 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-12
Perfect score: 3066
Sequence: 1 GGAACAGCTATGACCATGA.....GGTACCCAAATTCGCCCTATA 3066

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3066	100.0	3066	4	US-08-892-695-12
2	2813	91.7	2813	3	US-09-689-255C-3
3	2813	91.7	2813	4	US-09-949-016-531
4	2813	91.7	3629	4	US-09-949-016-4308
5	1018.4	33.2	130563	4	US-09-949-016-12273
6	1018.4	33.2	131379	4	US-09-949-016-16050
C 7	464.4	15.1	601	4	US-09-949-016-28396
C 8	464.4	15.1	601	4	US-09-949-016-153520
C 9	280	9.1	601	4	US-09-949-016-153435
C 10	238	7.8	297	4	US-09-513-999C-14399
C 11	236.8	7.7	72128	4	US-09-949-016-16018
C 12	234	7.6	42053	4	US-09-949-016-15924
C 13	232.2	7.6	137753	4	US-09-949-016-17404
C 14	231.8	7.6	601	4	US-09-949-016-201956
C 15	231	7.5	101011	4	US-09-949-016-16933
C 16	230.8	7.5	33712	4	US-09-949-016-15793
C 17	230.8	7.5	63930	4	US-09-949-016-12270
C 18	230.8	7.5	109690	4	US-09-949-016-13525
C 19	230.8	7.5	670689	4	US-09-949-016-12505
C 20	230.8	7.5	670690	4	US-09-949-016-14207
C 21	230.2	7.5	601	4	US-09-949-016-119390
C 22	230.2	7.5	601	4	US-09-949-016-119392
C 23	230	7.5	14345	4	US-09-949-016-15449
C 24	230	7.5	39243	4	US-09-949-016-12316
C 25	230	7.5	39243	4	US-09-949-016-15443
C 26	229.8	7.5	152132	4	US-09-949-016-13845
C 27	229.8	7.5	152145	4	US-09-949-016-12371

C 28	229.8	7.5	194937	4	US-09-949-016-17032	Sequence 17032, A
C 29	229.8	7.5	194937	4	US-09-949-016-17033	Sequence 17033, A
C 30	229.6	7.5	282	1	US-08-133-629-8	Sequence 8, Appli
C 31	229.6	7.5	601	4	US-09-949-016-131748	Sequence 131748, A
C 32	229.6	7.5	107085	4	US-09-949-016-13157	Sequence 13157, A
C 33	229.2	7.5	58593	4	US-09-949-016-12232	Sequence 12232, A
C 34	229	7.5	51403	4	US-09-949-016-15057	Sequence 15057, A
C 35	228.8	7.5	125672	4	US-09-949-016-16956	Sequence 16956, A
C 36	228.6	7.5	601	4	US-09-949-016-119391	Sequence 119391, A
C 37	228.6	7.5	601	4	US-09-949-016-138463	Sequence 138463, A
C 38	228.6	7.5	22634	4	US-09-949-016-16686	Sequence 16686, A
C 39	228.6	7.5	62804	3	US-09-800-960-3	Sequence 3, Appli
C 40	228.6	7.5	62804	4	US-10-096-960-3	Sequence 3, Appli
C 41	228.4	7.4	601	4	US-09-949-016-135308	Sequence 135308, A
C 42	228.4	7.4	26115	4	US-09-949-016-13599	Sequence 13599, A
C 43	228.4	7.4	107421	4	US-09-949-016-15532	Sequence 15532, A
C 44	228.2	7.4	97423	4	US-09-949-016-12742	Sequence 12742, A
C 45	228.2	7.4	97424	4	US-09-949-016-15576	Sequence 15576, A

ALIGNMENTS

RESULT 1
US-08-892-695-12
; Sequence 12, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A
; EARLIER FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 3066
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:1b1
US-08-892-695-12

Query Match	100.0%;	Score	3066;	DB	4;	Length	3066;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	3066;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	GGAACAGCTATGACCATGATTAGCCAAAGCTCGAAATTAACCCCTCACTAAAGGGAACAA	60				
Db	1	GGAACAGCTATGACCATGATTAGCCAAAGCTCGAAATTAACCCCTCACTAAAGGGAACAA	60				
QY	61	AAGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAG	120				
Db	61	AAGCTGGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGGATCCCCGGGCTGCAG	120				
QY	121	GAATTCGGCAGCGGCTCCACCGAGCCAGCCAGCGACTGGGCGAGCGACTGGAGAGCCAG	180				
Db	121	GAATTCGGCAGCGGCTCCACCGAGCCAGCCAGCGACTGGGCGAGCGACTGGAGAGCCAG	180				
QY	181	GACCTGTGCGAGGAGCAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCG	240				
Db	181	GACCTGTGCGAGGAGCAGCTCCGGGTGACACGAGGGGACTGAAGATATCTCCACAGGGGCG	240				

Qy 241 TCACGAGGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAT 300
Db |||||
Qy 241 TCACGAGGCAATGGGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACCAAGAGAAT 300
Db |||||
Qy 301 GAACGAGAGCAGAGACTTACGAGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Db |||||
Qy 301 GAACGAGAGCAGAGACTTACGAGACAAACGCGTCTGCTCTGAACGGGGTTCCAGTGGTG 360
Db |||||
Qy 361 GTGTGCGACCCACACAGTTTCAGCACTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGACG 420
Db |||||
Qy 361 GTGTGCGACCCACACAGTTTCAGCACTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGACG 420
Db |||||
Qy 421 GATAATGTGGCCACTTCTTCCCGGAGACAAACGAGATAGTGTGTGGGGATGCGCAAC 480
Db |||||
Qy 421 GATAATGTGGCCACTTCTTCCCGGAGACAAACGAGATAGTGTGTGGGGATGCGCAAC 480
Db |||||
Qy 481 GGAAGGAATCTTGGGAAGAGGCCAAACCCGAGGCACAGCTGTCTAAATCTCGTTTTTTC 540
Db |||||
Qy 481 GGAAGGAATCTTGGGAAGAGGCCAAACCCGAGGCACAGCTGTCTAAATCTCGTTTTTTC 540
Db |||||
Qy 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTCTATCCCTT 600
Db |||||
Qy 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTCTATCCCTT 600
Db |||||
Qy 601 GGATCAGTGAAGCTTGTATGTCTCAGCTCCAAATAAGCTCCAGCGAAACAAAGACCCCAAGTGAG 660
Db |||||
Qy 601 GGATCAGTGAAGCTTGTATGTCTCAGCTCCAAATAAGCTCCAGCGAAACAAAGACCCCAAGTGAG 660
Db |||||
Qy 661 AGCTGGACACTTCCGCTGGCAGCTGGACCCGGGGCAGACACAGATAAAACCCCAAGGCAC 720
Db |||||
Qy 661 AGCTGGACACTTCCGCTGGCAGCTGGACCCGGGGCAGACACAGATAAAACCCCAAGGCAC 720
Db |||||
Qy 721 GCGCGGCGCCAAAGACAGGTCTCTCTGCGCCAGGATCCACGCTTCTCCACCTGAG 780
Db |||||
Qy 721 GCGCGGCGCCAAAGACAGGTCTCTCTGCGCCAGGATCCACGCTTCTCCACCTGAG 780
Db |||||
Qy 781 ACAGGGGAGCAGGAGAGAGTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAA 840
Db |||||
Qy 781 ACAGGGGAGCAGGAGAGAGTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAA 840
Db |||||
Qy 841 TTCTTCAAGCTGGAACAAGGACAGAAAGGTGCCAGTGAACGCCAACAGGAAGCCAAAG 900
Db |||||
Qy 841 TTCTTCAAGCTGGAACAAGGACAGAAAGGTGCCAGTGAACGCCAACAGGAAGCCAAAG 900
Db |||||
Qy 901 AGGCGAGCATCAAGCAAGGTGGATGAGTTCCTGGCTTATCAGGCGATCCGATGAT 960
Db |||||
Qy 901 AGGCGAGCATCAAGCAAGGTGGATGAGTTCCTGGCTTATCAGGCGATCCGATGAT 960
Db |||||
Qy 961 GTCCCTGCAGGGAAGGACATAGTTGACGGCAAGGAAAGGACAAAGACTTTGGAACT 1020
Db |||||
Qy 961 GTCCCTGCAGGGAAGGACATAGTTGACGGCAAGGAAAGGACAAAGACTTTGGAACT 1020
Db |||||
Qy 1021 GCGGATTGCTCTGTCCCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCGCAAG 1080
Db |||||
Qy 1021 GCGGATTGCTCTGTCCCTGGGACCCAGAGGACTGGAGACTGCAAAAGGACGATTCGCAAG 1080
Db |||||
Qy 1081 GCAGCAGCTATAGCAGAGAATAATAATTCATCATGAGTTTCTTAAACCTCTGGTTCA 1140
Db |||||
Qy 1081 GCAGCAGCTATAGCAGAGAATAATAATTCATCATGAGTTTCTTAAACCTCTGGTTCA 1140
Db |||||
Qy 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCGAGACACGGGTGCTGAAAAGTCAACCCACC 1200
Db |||||
Qy 1141 CCTAAACAAAGCTGAAACAAAAAGGACCCGAGACACGGGTGCTGAAAAGTCAACCCACC 1200
Db |||||
Qy 1201 ACTTCAGCTGACCTTAAGTTCAGACAAAGCCAACTTTACATCCGAGGAGACCCAAAGGGCT 1260
Db |||||
Qy 1201 ACTTCAGCTGACCTTAAGTTCAGACAAAGCCAACTTTACATCCGAGGAGACCCAAAGGGCT 1260
Db |||||
Qy 1261 GGCAGGATCCAAAGGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTCAA 1320
Db |||||
Qy 1261 GGCAGGATCCAAAGGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTCAA 1320
Db |||||

Qy 1321 CCTGCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db |||||
Qy 1321 CCTGCGAAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTG 1380
Db |||||
Qy 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAAATGTGGTG 1440
Db |||||
Qy 1381 TTTTGGAAAAAGTCAGTTAAAGAGGACTCAGTCCCCACAGGTGCGGAGGAAATGTGGTG 1440
Db |||||
Qy 1441 TGTGAGTCACAGTAGAGATTATAAGTCCAAGGAAGTAGAATCAGCTCTTACAAAAGTGT 1500
Db |||||
Qy 1441 TGTGAGTCACAGTAGAGATTATAAGTCCAAGGAAGTAGAATCAGCTCTTACAAAAGTGT 1500
Db |||||
Qy 1501 GACCTCAACGAAGGAGATGTCACCTGAAACCCACAGAAGCGAAATCTCAAAAGAGAAGAA 1560
Db |||||
Qy 1501 GACCTCAACGAAGGAGATGTCACCTGAAACCCACAGAAGCGAAATCTCAAAAGAGAAGAA 1560
Db |||||
Qy 1561 AGCAAAACCAAGAACCTCTCTGATGGCGTTTCTCAGACAAATGTCAGTGAAGGGGATGGA 1620
Db |||||
Qy 1561 AGCAAAACCAAGAACCTCTCTGATGGCGTTTCTCAGACAAATGTCAGTGAAGGGGATGGA 1620
Db |||||
Qy 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAAATCATCAGACTCC 1680
Db |||||
Qy 1621 GGGATCACCCACTCAGAAAGAAATAAATGGGAAAGACTCCAGCTGCCAAAATCATCAGACTCC 1680
Db |||||
Qy 1681 ACAGAAAAGACTATCAACCCGCCAGAGCTTGAACCAAACAGGAGCACACAGAAAGGGTAAA 1740
Db |||||
Qy 1681 ACAGAAAAGACTATCAACCCGCCAGAGCTTGAACCAAACAGGAGCACACAGAAAGGGTAAA 1740
Db |||||
Qy 1741 GAGGGCTCTCGAAGGACAAAGAGTCAAGAGCCGAGATGAACAGACAGAGAGCAACAG 1800
Db |||||
Qy 1741 GAGGGCTCTCGAAGGACAAAGAGTCAAGAGCCGAGATGAACAGACAGAGAGCAACAG 1800
Db |||||
Qy 1801 CAGGAAGCCAAAGAAACCCAGCCAGTGCACAGAGCAGGCCACGGTGGACACGAACTCACTG 1860
Db |||||
Qy 1801 CAGGAAGCCAAAGAAACCCAGCCAGTGCACAGAGCAGGCCACGGTGGACACGAACTCACTG 1860
Db |||||
Qy 1861 CAGAAATGGGACAAAGCTTCCAAAAGAGACTTGAGAAAGCGGAGCAGTCCCTTGGGGGCTTC 1920
Db |||||
Qy 1861 CAGAAATGGGACAAAGCTTCCAAAAGAGACTTGAGAAAGCGGAGCAGTCCCTTGGGGGCTTC 1920
Db |||||
Qy 1921 TTTAAAGGCTGGGACCAAAAGCGGATGTTGGATGCTCAAGTGCAAAAAGAGAGTATCC 1980
Db |||||
Qy 1921 TTTAAAGGCTGGGACCAAAAGCGGATGTTGGATGCTCAAGTGCAAAAAGAGAGTATCC 1980
Db |||||
Qy 1981 ATCGGACAGTTGGCAAAACCCAGTAAACAAATCAGCACCGTTCGCCACAGGTTCTCTG 2040
Db |||||
Qy 1981 ATCGGACAGTTGGCAAAACCCAGTAAACAAATCAGCACCGTTCGCCACAGGTTCTCTG 2040
Db |||||
Qy 2041 CCACCAAGATGTGTCTCTTACT 2100
Db |||||
Qy 2041 CCACCAAGATGTGTCTCTTACT 2100
Db |||||
Qy 2101 TTCTGATGGCAGCAAAATGAAATTCCTCTAGAAATTAAGCCCGAGCTGTGTATATGTA 2160
Db |||||
Qy 2101 TTCTGATGGCAGCAAAATGAAATTCCTCTAGAAATTAAGCCCGAGCTGTGTATATGTA 2160
Db |||||
Qy 2161 GGTGTATTTATTCAGTCTCTGGTCCAGTCTTTCTGCAAAATAACAGTAAAGATGGTTTA 2220
Db |||||
Qy 2161 GGTGTATTTATTCAGTCTCTGGTCCAGTCTTTCTGCAAAATAACAGTAAAGATGGTTTA 2220
Db |||||
Qy 2221 GCAGGTCACTAGTGGGTGAGAGAGTTCGATGATCAACAGAGGAAAGAGGAGGGAATA 2280
Db |||||
Qy 2221 GCAGGTCACTAGTGGGTGAGAGAGTTCGATGATCAACAGAGGAAAGAGGAGGGAATA 2280
Db |||||
Qy 2281 GAGGAATGTGTTCGGGTAAAGTGAATAATGGCAGTGGTGGCCGGGGCTGGTGGCTCTC 2340
Db |||||
Qy 2281 GAGGAATGTGTTCGGGTAAAGTGAATAATGGCAGTGGTGGCCGGGGCTGGTGGCTCTC 2340
Db |||||
Qy 2341 GCCTGTAATCTCAGCACTTTGGGAGGCGGAGGAGTGGATCAGCTCAGGTCAGGAGTTC 2400
Db |||||
Qy 2341 GCCTGTAATCTCAGCACTTTGGGAGGCGGAGGAGTGGATCAGCTCAGGTCAGGAGTTC 2400
Db |||||
Qy 2401 AAGACTAGCCTGGCCAAACATCATGAAACCCCGCTCTCTACTAAATAATCAAAAAATTAGCCA 2460
Db |||||

Db 1081 AAGTCAGACAAAGCCAACTTTACATCCAGGAGACCACAAAGGGCTGGCAAGAAATTCACAA 1140
Qy 1276 GGATGCAACCATCGGGGCACACACAGTCCGTGACCAACCCCTGAACTCGGAAGGAAGGC 1335
Db 1141 GGATGCAACCATCGGGGCACACACAGTCCGTGACCAACCCCTGAACTCGGAAGGAAGGC 1200
Qy 1336 ACCAAGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db 1201 ACCAAGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGTGTGTGAGTCAACCAAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCAAGGTCGGAGGAGAAATGTGTGTGTGAGTCAACCAAGTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCCTTCAAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAGTAGAATCAGCCCTTCAAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1575
Db 1381 GATCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCA 1635
Db 1441 TCTCTGATGGCGTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCTCA 1500
Qy 1636 GAAGAAATAAGTGGAAAGACTCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1695
Db 1501 GAAGAAATAAGTGGAAAGACTCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAACCAACAGAGAGCACACAGAGGGTAAAGGGCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAACCAACAGAGAGCACACAGAGGGTAAAGGGCTCTCGAAG 1620
Qy 1756 GACAAGAGTCAGCAGCGGAGATGAAACAAGCAGAGAGCAACAAGAGGAGCAACAGAA 1815
Db 1621 GACAAGAGTCAGCAGCGGAGATGAAACAAGCAGAGAGCAACAAGAGGAGCAACAGAA 1680
Qy 1816 CCAGCCCATGTCACAGAGAGCGCCACCGTGACACGAACCTCAGTGCAGATGGGGACAAG 1875
Db 1681 CCAGCCCATGTCACAGAGAGCGCCACCGTGACACGAACCTCAGTGCAGATGGGGACAAG 1740
Qy 1876 CTCACAAAGAGACTGTGAGAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1935
Db 1741 CTCACAAAGAGACTGTGAGAGCGCAGCAGTCCCTTGGGGCTTCTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCACACAGACCCAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAACCCCAAGTAAACAAATCAGACACGGTTCACACAGGTTCTCTGCCACCAAGATGTT 2055
Db 1861 AAACCCCAAGTAAACAAATCAGACACGGTTCACACAGGTTCTCTGCCACCAAGATGTT 1920
Qy 2056 CTCCTTACTCCATCTCTCCCAACACACGCTCCATGTATATATTTCTGTATGCCACGCA 2115
Db 1921 CTCCTTACTCCATCTCTCCCAACACACGCTCCATGTATATATTTCTGTATGCCACGCA 1980
Qy 2116 AATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTAGGTTGATATTATTACG 2175
Db 1981 AATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTAGGTTGATATTATTACG 2040
Qy 2176 TCTCTGCTCAGCTCTTTCTGGCAAAATACAGTAAAGATGTTTACAGGTCACCTAGTT 2235
Db 2041 TCTCTGCTCAGCTCTTTCTGGCAAAATACAGTAAAGATGTTTACAGGTCACCTAGTT 2100
Qy 2236 GGGTCAGAAGAGTCAGTATCACAAGCAGGAGGAGGGAATAGAGGAATGTTCGG 2295
Db 2101 GGGTCAGAAGAGTCAGTATCACAAGCAGGAGGAGGGAATAGAGGAATGTTCGG 2160
Qy 2296 GTTAAAGTATGAAATGSCAGTGTGGCCGGCGGTGGTCTCGCTGCTGTAATCTCAGC 2355

Db 2161 GTTAAAGTATGAAATGCGAGTGGTGGCCGGCGGTGGTGGCTCTCGCTCTGTAATCTCAGC 2220
Qy 2356 ACTTTGGAGCGCGAGCGAGTGGATCACTGAGGTTCAGAGTTCAGAGCTAGCCTGGCC 2415
Db 2221 ACTTTGGAGCGCGAGCGAGTGGATCACTGAGGTTCAGAGTTCAGAGCTAGCCTGGCC 2280
Qy 2416 AACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCAGGCGATGGTGGCACAC 2475
Db 2281 AACATCATGAAACCCCGTCTCTACTAAAAATACAAAAATTAGCAGGCGATGGTGGCACAC 2340
Qy 2476 ACCTGTAGTCCAGCTACTCGGAGCGCCAAACGACAGAAACCGCTTGTATCCAGGAGGTG 2535
Db 2341 ACCTGTAGTCCAGCTACTCGGAGCGCCAAACGACAGAAACCGCTTGTATCCAGGAGGTG 2400
Qy 2536 GAGTTTGCAGTGGAGCGCAAGTTGCAACATTCGACTCCACCTGGGCGACAGAGCAAGATT 2595
Db 2401 GAGTTTGCAGTGGAGCGCAAGTTGCAACATTCGACTCCACCTGGGCGACAGAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAAGAGAAATGCTGTAGAAGGAA 2655
Db 2461 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAAGAGAAATGCTGTAGAAGGAA 2520
Qy 2656 TTAAGCGTTGTAGTAAACGCTGCTCATCTCTTAAGCTTGAAGAGGAGGACGAAAATCC 2715
Db 2521 TTAAGCGTTGTAGTAAACGCTGCTCATCTCTTAAGCTTGAAGAGGAGGACGAAAATCC 2580
Qy 2716 ATTGTGTTAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTGGTGGTTGCCAAT 2775
Db 2581 ATTGTGTTAAATTCACATCTCAAGGAGGAGAACCCGGCTGTGTGGTGGTTGCCAAT 2640
Qy 2776 TTCTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGCTGTGTTTTCAAAT 2835
Db 2641 TTCTAGAACCGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGCTGTGTTTTCAAAT 2700
Qy 2836 AGGTCTCTGTAAATTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAAAATGAT 2895
Db 2701 AGGTCTCTGTAAATTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAAAATGAT 2760
Qy 2896 TTGGGAAACAAATGCTTTTGGGCTCAGTGACAAACGCGCAAGATTACACTT 2948
Db 2761 TTGGGAAACAAATGCTTTTGGGCTCAGTGACAAACGCGCAAGATTACACTT 2813

RESULT 3

US-09-949-016-531
; Sequence 531, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-531

Query Match 91.7%; Score 2813; DB 4; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 136 CTCACCCGACAGCCAGGCGACTGGGCGACGCACTGGAGACCCAGGACCCCTGTGCAAGAG 195


```
QY 2356 ACTTTGGAGGCCGAGCAGGTGGATCACTGAGGTGAGGAGTTCAAGACTAGCCTGGCC 2415
Db |||
QY 2221 ACTTTGGAGGCCGAGCAGGTGGATCACTGAGGTGAGGAGTTCAAGACTAGCCTGGCC 2280
Db |||
QY 2416 AACATCATGAACCCCGTCTCTACTATAAATACAAAAATAGCCAGGCATGGTGGCACAC 2475
Db |||
QY 2281 AACATCATGAACCCCGTCTCTACTATAAATACAAAAATAGCCAGGCATGGTGGCACAC 2340
QY 2476 ACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2535
Db |||
QY 2341 ACCTGTAGTCCAGTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2400
QY 2536 GAGGTTGAGTGAGCCGAAAGTTGACCAATTGCACTCCACCTCGGCGACAGCAAGATT 2595
Db |||
QY 2401 GAGGTTGAGTGAGCCGAAAGTTGACCAATTGCACTCCACCTCGGCGACAGCAAGATT 2460
QY 2596 CTATCAAAAAAAGGAGGTGGCAAGTAAGTATAGAGAAATGCTGCTAGAGGAA 2655
Db |||
QY 2461 CTATCAAAAAAAGGAGGTGGCAAGTAAGTATAGAGAAATGCTGCTAGAGGAA 2520
QY 2656 TTAAGCGTTGTAGTAAGCGGTGCTCATCTTAAGCTTGAAGAGGAGACGAAATCC 2715
Db |||
QY 2521 TTAAGCGTTGTAGTAAGCGGTGCTCATCTTAAGCTTGAAGAGGAGACGAAATCC 2580
QY 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGGTGGTTGCCAAT 2775
Db |||
QY 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGGTGGTTGCCAAT 2640
QY 2776 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2835
Db |||
QY 2641 TTCTTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAAAT 2700
QY 2836 AGGTCCTTGTAGTATGATGAGGAGGAAAGATTGACTGGGAGGGCTTAAATGAT 2895
Db |||
QY 2701 AGGTCCTTGTAGTATGATGAGGAGGAAAGATTGACTGGGAGGGCTTAAATGAT 2760
QY 2896 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACCGCAAGAAATTACAACTT 2948
Db |||
QY 2761 TTGGGAAAAAATTTGCTTTTGGGCTCAGTGACAAACCGCAAGAAATTACAACTT 2813

RESULT 4
US-09-949-016-4308
; Sequence 4308, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4308
; LENGTH: 3629
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4308

Query Match 91.7%; Score 2813; DB 4; Length 3629;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2827; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

QY 136 CTCACCGACAGCCAGGCACTGGGCGAGCAGCACTGGAGACCCAGGACCCCTGTGCGAGG 195
Db |||
```

```
Db 1 CTCACCGACAGCCAGGCACTGGGCGAGCAGCACTGGAGACCCAGGACCCCTGTGCGAGG 60
QY 196 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 255
Db |||
QY 61 CAGCTCCGGGTGACACGAGGGGACTGAAGATACTCCCAAGGGGCTCAGCAGGAGCAATG 120
QY 256 GGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACAAGAGAAATGAACAGAGAGCAG 315
Db |||
QY 121 GGTAAACCAATGAGTGTTCCTCCAAAGAGTTGAAGACAAGAGAAATGAACAGAGAGCAG 180
QY 316 ACTTACCGACAGCAACCGTCTGCTCGAACGGGTTCCAGTGTGTGTCGACCCACACA 375
Db |||
QY 181 ACTTACCGACAGCAACCGTCTGCTCGAACGGGTTCCAGTGTGTGTCGACCCACACA 240
QY 376 GTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGAAGTAATGTGGCCACT 435
Db |||
QY 241 GTTCAGCACTTAGAGGAAGTTCGACTTGGGAATAAGTGTCAAGAAGTAATGTGGCCACT 300
QY 436 TCTTCCCGGAGACAAACCGGAGATAAGTGTGTGGGATGCCAACGGAAAGAAATCTTTGGG 495
Db |||
QY 301 TCTTCCCGGAGACAAACCGGAGATAAGTGTGTGGGATGCCAACGGAAAGAAATCTTTGGG 360
QY 496 AAAGAGCCCAAAACCGGAGACCAAGCCGAGACTCATCCCTTGGATCAGTGAAGCTT 615
Db |||
QY 421 CTTGTACAGAGACGTAACCGGAGACCAAGCCGAGACTCATCCCTTGGATCAGTGAAGCTT 480
QY 616 GATGTACAGCTTCAATAAAGCTCCAGCGAAACAAAGACCAAGTGAAGTGAAGTGAAGTGA 675
Db |||
QY 481 GATGTACAGCTTCAATAAAGCTCCAGCGAAACAAAGACCAAGTGAAGTGAAGTGAAGTGA 540
QY 676 GTGGCAGCTGGAACCGGGGAGGACACAGATAAACCCCGGCGACGCCCGGCGCCAGAGAC 735
Db |||
QY 541 GTGGCAGCTGGAACCGGGGAGGACACAGATAAACCCCGGCGACGCCCGGCGCCAGAGAC 600
QY 736 AAGTGTCTCTCTGCGCGCAGGGATCCCAACGCTTCTCCCACTGAGACAGGGGGAGCAGGA 795
Db |||
QY 601 AAGTGTCTCTCTGCGCGCAGGGATCCCAACGCTTCTCCCACTGAGACAGGGGGAGCAGGA 660
QY 796 GGAGAAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 855
Db |||
QY 661 GGAGAAGCTCCCTCCAAAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 720
QY 856 AAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAGCCCAAGAGGCGAGAGCATCAA 915
Db |||
QY 721 AAGGACAGGAAAGGTGCCAGGTGACAGCCCAAGAGCCCAAGAGGCGAGAGCATCAA 780
QY 916 GACAAGTGTGATGAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCGAGGGAAG 975
Db |||
QY 781 GACAAGTGTGATGAGGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCCTGCGAGGGAAG 840
QY 976 GACATAGTTGACGCAAGGAAAAAAGAGGACAAGAACTTGGAACTGGCGATGCTCTGTC 1035
Db |||
QY 841 GACATAGTTGACGCAAGGAAAAAAGAGGACAAGAACTTGGAACTGGCGATGCTCTGTC 900
QY 1036 CTTGGGACCCAGAGGACTGGGACACTGCAAGAGCAATTCAGGAGGAGCAGCTATAGCA 1095
Db |||
QY 901 CTTGGGACCCAGAGGACTGGGACACTGCAAGAGCAATTCAGGAGGAGCAGCTATAGCA 960
QY 1096 GAGAATAATAATTCATCATGAGTTTCTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1155
Db |||
QY 961 GAGAATAATAATTCATCATGAGTTTCTTTAAAACTCTGGTTTCACTTAAACAAAGCTGAA 1020
QY 1156 AAAAAAAGGACCCAGAGGACACGGGTGCTGAAAAAGTCAACCACTTCAGCTGACCTT 1215
Db |||
QY 1021 AAAAAAAGGACCCAGAGGACACGGGTGCTGAAAAAGTCAACCACTTCAGCTGACCTT 1080
QY 1216 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCCAAGGGCTGGCAAGAAATTCAAA 1275
Db |||
QY 1081 AAGTCAGACAAAGCCAACTTTACATCCAGAGAGCCCAAGGGCTGGCAAGAAATTCAAA 1140
```


Qy	1929	CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC	1988
Db	127543	CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC	127602
Qy	1989	AGTTGGCAAAACCCAAAGTAACAATCAGCAGCGTTCACACAGGTTCTCTGCGACCAAG	2048
Db	127603	AGTTGGCAAAACCCAAAGTAACAATCAGCAGCGTTCACACAGGTTCTCTGCGACCAAG	127662
Qy	2049	ATGTGTTCTCCTTACTCCATCTCTCCGCAACAGCGTCCATGTATATATCTTCTGATG	2108
Db	127663	ATGTGTTCTCCTTACTCCATCTCTCCGCAACAGCGTCCATGTATATATCTTCTGATG	127722
Qy	2109	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	2168
Db	127723	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	127782
Qy	2169	ATTTACGCTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGGTTTAGAGTCA	2228
Db	127783	ATTTACGCTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGGTTTAGAGTCA	127842
Qy	2229	CCTAGTTGGTTCAGAGTGCATGATCACCAGCAGGAAGGAGGGAATAGAGGAATG	2288
Db	127843	CCTAGTTGGTTCAGAGTGCATGATCACCAGCAGGAAGGAGGGAATAGAGGAATG	127902
Qy	2289	TGTTCCGGTTAAAGTGAATGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	2348
Db	127903	TGTTCCGGTTAAAGTGAATGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	127962
Qy	2349	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	2408
Db	127963	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	128022
Qy	2409	CCTGGCAACATCATGAAATCCCGTCTCTACTTAAATAACAAAATTAAGCCAGCATGGT	2468
Db	128023	CCTGGCAACATCATGAAATCCCGTCTCTACTTAAATAACAAAATTAAGCCAGCATGGT	128082
Qy	2469	GGCACACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGATCCCA	2528
Db	128083	GGCACACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGATCCCA	128142
Qy	2529	GGAGTGGAGGTTGCAGTAGCCGAAGTTGCAACATTCGACTCCACCTCGGGCGACAGAG	2588
Db	128143	GGAGTGGAGGTTGCAGTAGCCGAAGTTGCAACATTCGACTCCACCTCGGGCGACAGAG	128202
Qy	2589	CAAGATTTCTATC-NAAAAAAAAAAAGGCGAGTGCAGTAAGTTATAGAGAGAAATGCTGCT	2647
Db	128203	CAAGATTTCTATC-NAAAAAAAAAAAGGCGAGTGCAGTAAGTTATAGAGAGAAATGCTGCT	128262
Qy	2648	AGAGGAATTAAGCGTTGTAGTAAACGGTCTCATCTCTAAGCTTGAAGAGGGAGAC	2707
Db	128263	AGAGGAATTAAGCGTTGTAGTAAACGGTCTCATCTCTAAGCTTGAAGAGGGAGAC	128322
Qy	2708	GAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG	2767
Db	128323	GAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG	128382
Qy	2768	TTGCCAATTTCTAGAACGGAATGTGGGTATAGAAAAGGAATGAATAAGCGTTGTT	2827
Db	128383	TTGCCAATTTCTAGAACGGAATGTGGGTATAGAAAAGGAATGAATAAGCGTTGTT	128442
Qy	2828	TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAAGATTGACTGGGAGGGCTT	2887
Db	128443	TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAAGATTGACTGGGAGGGCTT	128502
Qy	2888	AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTTACAAT	2947
Db	128503	AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTTACAAT	128562
Qy	2948	AAAAAAAAAAAAAAAAAAAAA 2967	
Db	128563	AAAAAAAAAAAAAAAAAAAAA 128582	

RESULT 6			
US-09-949-016-16050			
; Sequence 16050, Application US/09949016			
; Patent No. 6812339			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL001307			
; CURRENT APPLICATION NUMBER: US/09/949,016			
; CURRENT FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-04-14			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-20			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-10-03			
; PRIOR FILING DATE: 2000-09-08			
; NUMBER OF SEQ ID NOS: 207012			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 16050			
; LENGTH: 131379			
; TYPE: DNA			
; ORGANISM: Human			
US-09-949-016-16050			
Query Match			
Best Local Similarity 33.2%; Score 1018.4; DB 4; Length 131379;			
Matches 1033; Conservative 0; Mismatches 6; Indels 1; Gaps 1;			
Qy	1929	CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC	1988
Db	127543	CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCAGTATCCATCGGACC	127602
Qy	1989	AGTTGGCAAAACCCAAAGTAACAATCAGCAGCGTTCACACAGGTTCTCTGCGACCAAG	2048
Db	127603	AGTTGGCAAAACCCAAAGTAACAATCAGCAGCGTTCACACAGGTTCTCTGCGACCAAG	127662
Qy	2049	ATGTGTTCTCCTTACTCCATCTCTCCGCAACAGCGTCCATGTATATATCTTCTGATG	2108
Db	127663	ATGTGTTCTCCTTACTCCATCTCTCCGCAACAGCGTCCATGTATATATCTTCTGATG	127722
Qy	2109	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	2168
Db	127723	GCACGAAATGAATTTCTGCTAGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	127782
Qy	2169	ATTTACGCTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGGTTTAGAGTCA	2228
Db	127783	ATTTACGCTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGGTTTAGAGTCA	127842
Qy	2229	CCTAGTTGGTTCAGAGTGCATGATCACCAGCAGGAAGGAGGGAATAGAGGAATG	2288
Db	127843	CCTAGTTGGTTCAGAGTGCATGATCACCAGCAGGAAGGAGGGAATAGAGGAATG	127902
Qy	2289	TGTTCCGGTTAAAGTGAATGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	2348
Db	127903	TGTTCCGGTTAAAGTGAATGAAATTAAGCCCGAGCTGTTGTATATATTGAGTGTATT	127962
Qy	2349	TCTCAGCACTTTGGGAGCCGAGCGAGTGGATCACCTGAGGTCAGAGTGTCAAGACTAG	2408
Db	127963	TCTCAGCACTTTGGGAGCCGAGCGAGTGGATCACCTGAGGTCAGAGTGTCAAGACTAG	127662
Qy	2409	CCTGGCAACATCATGAAATCCCGTCTCTACTTAAATAACAAAATTAAGCCAGCATGGT	2468
Db	128023	CCTGGCAACATCATGAAATCCCGTCTCTACTTAAATAACAAAATTAAGCCAGCATGGT	127662
Qy	2469	GGCACACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGATCCCA	2528
Db	128083	GGCACACACCTGTAGTCCAGCTACTCGGGAGCCCAACGACGAGAACCGCTTGATCCCA	127722
Qy	2529	GGAGTGGAGGTTGCAGTAGCCGAAGTTGCAACATTCGACTCCACCTCGGGCGACAGAG	2588
Db	128143	GGAGTGGAGGTTGCAGTAGCCGAAGTTGCAACATTCGACTCCACCTCGGGCGACAGAG	127782
Qy	2589	CAAGATTTCTATC-NAAAAAAAAAAAGGCGAGTGCAGTAAGTTATAGAGAGAAATGCTGCT	2647
Db	128203	CAAGATTTCTATC-NAAAAAAAAAAAGGCGAGTGCAGTAAGTTATAGAGAGAAATGCTGCT	127842
Qy	2648	AGAGGAATTAAGCGTTGTAGTAAACGGTCTCATCTCTAAGCTTGAAGAGGGAGAC	2707
Db	128263	AGAGGAATTAAGCGTTGTAGTAAACGGTCTCATCTCTAAGCTTGAAGAGGGAGAC	127902
Qy	2708	GAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG	2767
Db	128323	GAAATCCATTTGTTTAAATTCATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGG	128382
Qy	2768	TTGCCAATTTCTAGAACGGAATGTGGGTATAGAAAAGGAATGAATAAGCGTTGTT	2827
Db	128383	TTGCCAATTTCTAGAACGGAATGTGGGTATAGAAAAGGAATGAATAAGCGTTGTT	128442
Qy	2828	TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAAGATTGACTGGGAGGGCTT	2887
Db	128443	TTTCAAAATAGGGTCTTGTAAAGTTATTGATCAGAGGGGAAAAGATTGACTGGGAGGGCTT	128502
Qy	2888	AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTTACAAT	2947
Db	128503	AAAAATGATTTGGGAAAAAATTTGCTTTTGGAGCTCAGTGACAAACCGCAAGATTTACAAT	128562
Qy	2948	AAAAAAAAAAAAAAAAAAAAA 2967	
Db	128563	AAAAAAAAAAAAAAAAAAAAA 128582	


```
Db 108 TGTTCGGGTTAAGTGAATGCAAAATGCGAGTGTGGCCGGGCTGTGGCTCTCGCCTGTA 49
Qy 2349 TCTCAGCACTTTGGGAGGCGGAGCGAGTGGATCACCTGAGGTCAGGA 2396
Db 48 TCTCAGCACTTTGGGAGGCGGAGCGAGTGGATCACCTGAGGTCAGGA 1

RESULT 9
US-09-949-016-153435/c
; Sequence 153435, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 153435
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-153435

Query Match 9.1%; Score 280; DB 4; Length 601;
Best Local Similarity 100.0%; Pred. No. 4.9e-64;
Matches 280; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 452
Db 280 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACTTCTTCCCGGAGACAAC 221

Qy 453 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAGAGGCCAACCCGA 512
Db 220 GGAGATAAGTCTGTTGGGATGCCAAGCGAAGAAATCTTGGGAAGAGGCCAACCCGA 161

Qy 513 GGACACGAGCTGCTAAATCTGTTTTTTCTTGATCTCTCTCGGCTGTACCAGACGTAC 572
Db 160 GGACACGAGCTGCTAAATCTGTTTTTTCTTGATCTCTCTCGGCTGTACCAGACGTAC 101

Qy 573 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGATGTCACTCCAATAA 632
Db 100 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGATGTCACTCCAATAA 41

Qy 633 AGCTCCAGCGCAACAAAGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 672
Db 40 AGCTCCAGCGCAACAAAGACCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1

RESULT 10
US-09-513-999C-14399
; Sequence 14399, Application US/0951399C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.Y.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59 US2,REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-02-26
```

```
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14399
; LENGTH: 297
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 223
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 237
; OTHER INFORMATION: m=a or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 267
; OTHER INFORMATION: k=g or t
; OTHER INFORMATION: k=g or t
US-09-513-999C-14399

Query Match 7.8%; Score 238; DB 4; Length 297;
Best Local Similarity 98.0%; Pred. No. 4.5e-53;
Matches 249; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

Qy 2715 CATTGTGTTAAATTCATCTCAAGGAGGAGAACCCGGCTGTGTGGTGTGGTGTGGTGTGG 2774
Db 1 CATTGTGTTAAATTCATCTCAAGGAGGAGAACCCGGCTGTGTGGTGTGGTGTGGTGTGGTGTGG 60

Qy 2775 TTTCTAGAACGGAATCTGTGGGATAGAAAAGGAATGAATGAATGAATGAATGAATGAATGA 2834
Db 61 TTTCTAGAACGGAATCTGTGGGATAGAAAAGGAATGAATGAATGAATGAATGAATGAATGAATGA 120

Qy 2835 TAGGTCCTTGTAAATTTATTG-ATGAGAGGAGAAAAGATTGACTGGGAGGCTTAAATG 2893
Db 121 TAGGTCCTTGTAAATTTATTG-ATGAGAGGAGAAAAGATTGACTGGGAGGCTTAAATG 180

Qy 2894 ATTTGGAAAAACAATTGCTTTTGGAGCTCAGTGAACCGGCAAGATACAACTTAAAAA 2953
Db 181 ATTTGGAAAAACAATTGCTTTTGGAGCTCAGTGAACCGGCAAGATACAACTTAAAAA 240

Qy 2954 AAAAAAAAAAAAAA 2967
Db 241 AAAAAAAAAATAA 254

RESULT 11
US-09-949-016-16018/c
; Sequence 16018, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16018
; LENGTH: 72128
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16018

Query Match 7.7%; Score 236.8; DB 4; Length 72128;
Best Local Similarity 78.6%; Pred. No. 1.4e-51;
Matches 283; Conservative 0; Mismatches 77; Indels 0; Gaps 0;
```


FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 201956
LENGTH: 601
TYPE: DNA
ORGANISM: Human
US-09-949-016-201956

Query Match 7.6%; Score 231.8; DB 4; Length 601;

Best Local Similarity 79.8%; Pred. No. 2.8e-51;
Matches 272; Conservative 1; Mismatches 68; Indels 0; Gaps 0;

Qy	2305	TGAAATGGCAGTGGTGGCGGGGGTGGTGGTCTCGCTCTGCTGTAATCTCAGCACCTTTGGGA	2364
Db	22	TTAAATGGCACCAAAATCGGGCGCGTGGCTCACGCTGTATCCAGCACCTTTGGGA	81
Qy	2365	GGCGAGCGAGGTGGATCACCCTGAGGTGAGGTTCAAGACTAGCCTGGCCCAACATCATG	2424
Db	82	GGCGAGCGAGGAGATCATTGAGGTGAGGTTTGAGACCAAGCCTGGCCCAACATGATA	141
Qy	2425	AAACCCGCTCTACTATAAATAACAAAATAGCCAGGCATGGTGGCACACCTGTAGT	2484
Db	142	AAACCCCATCTCTACTATAAATAACAAAATATCCAGGCGTGGTGGCTTACCTGTAGT	201
Qy	2485	CCAGCTACTCGGAGCGCCCAACGACGAGAACCGCTTGTACCCAGAGAGTGGAGTTGCA	2544
Db	202	CCAGCTACTCAGGAGATGAGCGAGAGATGGCTTTGAACCTGGAGCGGAGTTGCA	261
Qy	2545	GTGAGCGGAAGTTGCACCATTTGCACTCCACCTGGGCGGACAGACAAAGATTCTATCAAAA	2604
Db	262	GTGAGCGGAGATCACATCCTGCACTCCAGCTGGTGGTGAAGAGCGAGACTCCCATCTCAA	321
Qy	2605	AAAAAGGCGAGTGGCAAGTAAAGTTATAGAGAGAAATGCTG	2645
Db	322	AAAAAAAATGGCAACCAAGTATTTTCAGAACTAGGCGAG	362

RESULT 15

US-09-949-016-16933/c
Sequence 16933, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16933
LENGTH: 101011
TYPE: DNA
ORGANISM: Human
US-09-949-016-16933

Query Match 7.5%; Score 231; DB 4; Length 101011;
Best Local Similarity 83.9%; Pred. No. 5.7e-50;

	Matches	261;	Conservative	0;	Mismatches	50;	Indels	0;	Gaps	0;
Qy	2304	ATCAAAATGSCACTGGTGGCGGGCGTGGTGGCTCTCGCTGTAATCTCAGCACCTTTGGG	2363							
Db	5779	ATAAAAAATGATTTTAGGGCTGGCGGGTGGCTCAGCGCTGTAATCCCAAGCACTTTGGG	5720							
Qy	2364	AGCCCGAGGCGAGTGGATCACCTGAGGTCAAGAGTTCAAGACTAGCCTGGCCCAACATCAT	2423							
Db	5719	AGGCTGAGGCGAGCGAATCACCTGAGGTCAAGAGTTCAAGACCAAGCTGGCCCAACATGAT	5660							
Qy	2424	GRAACCCCGTCTCTACTATAAATAACAAAATTAAGCCAGGCATGTTGGCACACACCTGTAG	2483							
Db	5659	GAACCCCTGTCTCTACTATAAATAACAAAATTAAGAGGGCTGTTGGCACACACCTATAA	5600							
Qy	2484	TCCAGCTACTCGGAGCGCCCAACGACGAGAACCGCTTGTACCCAGAGAGTGGAGTTGC	2543							
Db	5599	TTCCAGCTACTTGGAGGGCGGAGGAGAACCTTTGAACCCAGAGGTGGAGTTGC	5540							
Qy	2544	AGTGAGCGGAAGTTGCACCATTTGCATCTCCACCTGGGCGGACAGACAAAGATTCTATCAAA	2603							
Db	5539	AGTGAGCGGAGATCATGCCACTGCACTCCAGCCTGGGCAACAGAGTGAGACCCCTGTCTCA	5480							
Qy	2604	AAAAAAGGCA	2614							
Db	5479	AAAAACAACA	5469							

Search completed: July 3, 2005, 16:59:43

Job time : 487.521 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 1788.95 Seconds

(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-12

Perfect score: 3066

Sequence: 1 GGAAACAGCTATGACCATGA.....GGTACCAATGCGCCATATA 3066

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	3066	100.0	3066	8	US-08-731-499-12
2	2828.8	92.3	3418	20	US-10-723-860-6167
3	2813	91.7	2813	17	US-10-172-118-937
4	2813	91.7	2813	18	US-10-240-425-385
5	2813	91.7	2813	18	US-10-342-887-937
6	2813	91.7	2813	20	US-10-723-860-1812
7	1929.4	62.9	2020	21	US-10-491-213-85
					Sequence 12, Appl
					Sequence 6167, Ap
					Sequence 937, App
					Sequence 385, App
					Sequence 937, App
					Sequence 1812, Ap
					Sequence 85, Appl

8	1567	51.1	2105	9	US-09-939-825-15	Sequence 15, Appl
9	1000.6	32.6	127567	22	US-10-737-082-47	Sequence 47, Appl
10	1000.6	32.6	127567	22	US-10-765-790-47	Sequence 47, Appl
c 11	582.6	19.0	644	16	US-10-029-386-22674	Sequence 22674, A
c 12	398.6	13.0	478	9	US-09-783-590-10392	Sequence 10392, A
c 13	332.8	10.9	576	16	US-10-029-386-3914	Sequence 3914, Ap
c 14	328	10.7	372	21	US-10-800-322-327	Sequence 327, App
c 15	297	9.7	506	16	US-10-029-386-8953	Sequence 8953, Ap
c 16	265	8.6	265	9	US-09-864-761-28309	Sequence 28309, A
c 17	260.4	8.5	495	9	US-09-864-761-11735	Sequence 11735, A
c 18	247	8.1	349	13	US-10-040-739-889	Sequence 889, App
c 19	242.6	7.9	6096	14	US-10-012-6009-132	Sequence 132, App
c 20	236	7.7	181684	13	US-10-087-192-790	Sequence 790, App
c 21	233.8	7.6	58355	21	US-10-484-577-665	Sequence 665, App
c 22	233.8	7.6	186591	21	US-10-484-577-682	Sequence 682, App
c 23	233.8	7.6	208648	21	US-10-484-577-668	Sequence 668, App
c 24	233.8	7.6	208648	21	US-10-484-577-668	Sequence 668, App
c 25	233.8	7.6	208648	21	US-10-484-577-668	Sequence 668, App
c 26	232.4	7.6	227246	19	US-10-322-281-314	Sequence 314, App
c 27	231.8	7.6	627	13	US-10-027-632-250385	Sequence 250385, A
c 28	231.8	7.6	627	17	US-10-027-632-250385	Sequence 250385, A
c 29	231.6	7.6	24699	9	US-09-764-877-2419	Sequence 2419, Ap
c 30	231.6	7.6	24699	17	US-10-242-515-2419	Sequence 2419, Ap
c 31	231.6	7.6	414295	20	US-10-719-993-6876	Sequence 6876, Ap
c 32	231.2	7.5	794	13	US-10-027-632-171663	Sequence 171663, A
c 33	231.2	7.5	794	17	US-10-027-632-171663	Sequence 171663, A
c 34	231	7.5	22047	20	US-10-417-375-12	Sequence 12, Appl
c 35	231	7.5	22056	19	US-10-322-281-106	Sequence 106, App
c 36	230.8	7.5	28693	21	US-10-741-600-17761	Sequence 17761, A
c 37	230.8	7.5	54732	19	US-10-322-281-414	Sequence 414, App
c 38	230.6	7.5	383	9	US-09-960-352-10031	Sequence 10031, A
c 39	230.4	7.5	227931	17	US-10-085-117-274	Sequence 274, App
c 40	229.8	7.5	253861	19	US-10-741-601-5611	Sequence 5611, Ap
c 41	229.8	7.5	261817	13	US-10-087-192-2002	Sequence 2002, Ap
c 42	229.4	7.5	91760	13	US-10-087-192-844	Sequence 844, App
c 43	229.4	7.5	379652	21	US-10-481-613-71	Sequence 71, Appl
c 44	229.2	7.5	285020	13	US-10-087-192-1666	Sequence 1666, Ap
c 45	229	7.5	285	21	US-10-956-157-121436	Sequence 121436, A

ALIGNMENTS

RESULT 1

US-08-731-499-12
; Sequence 12, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROWMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435


```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3066 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..3066
; OTHER INFORMATION: /note="lbi"
US-08-731-499-12

Query Match      100.0%; Score 3066; DB 8; Length 3066;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3066; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGAACAGCTATGACCATGATTACGCCAAGCTCGAATTAACCTCTCACTAAAGGACAA 60
DB 1 GGAACAGCTATGACCATGATTACGCCAAGCTCGAATTAACCTCTCACTAAAGGACAA 60
QY 61 AAGCTGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGTGATCCCGCGGCTGCAG 120
DB 61 AAGCTGAGCTCCACCGCGGTGGCGCGCTCTAGAACTAGTGTGATCCCGCGGCTGCAG 120
QY 121 GAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGAGCAACGACTGGAGACCAG 180
DB 121 GAATTCGGCAGAGGCTCCACCGACAGCCAGGCACTGGGAGCAACGACTGGAGACCAG 180
QY 181 GACCTGTGAGGAGAGCTCCCGGTGACAGGAGGACTGAGATACCTCCACAGGGGC 240
DB 181 GACCTGTGAGGAGAGCTCCCGGTGACAGGAGGACTGAGATACCTCCACAGGGGC 240
QY 241 TCAGCAGGAGCAATGGGTAAACCAATGAGTGTCTCCCAAGAGAGTTGAAGCAAGAAAT 300
DB 241 TCAGCAGGAGCAATGGGTAAACCAATGAGTGTCTCCCAAGAGAGTTGAAGCAAGAAAT 300
QY 301 GAACCAAGCAGAGACTTTACCAAGGACAAACGCGTCTGCTGAAACCGGGTTCCAGTGGTG 360
DB 301 GAACCAAGCAGAGACTTTACCAAGGACAAACGCGTCTGCTGAAACCGGGTTCCAGTGGTG 360
QY 361 GTGTGACCCACACAGTTACGACTTAGAGGAGTTCGACTTGGGAATAAGTGTCAAGACG 420
DB 361 GTGTGACCCACACAGTTACGACTTAGAGGAGTTCGACTTGGGAATAAGTGTCAAGACG 420
QY 421 GATAATGTGGCCACTTCTTCCCGCAGACAAACGAGATAAGTGTGTGGGGATGCCAAC 480
DB 421 GATAATGTGGCCACTTCTTCCCGCAGACAAACGAGATAAGTGTGTGGGGATGCCAAC 480
QY 481 GGAAGAATCTTGGGAAGAGGCCAAACCCGAGGCAACGCTGCTAAATCTCGTTTTTTC 540
DB 481 GGAAGAATCTTGGGAAGAGGCCAAACCCGAGGCAACGCTGCTAAATCTCGTTTTTTC 540
QY 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTTCATCCCTT 600
DB 541 TTGATGCTCTCTCGGCTGTACAGGAGCTACCGGAGACCAAGCCGAGATTTCATCCCTT 600
QY 601 GGATCAGTGAAGCTTGTGATGTGAGTCTCAATAAAGCTCCAGGCAACAAAGCCCAAGTGAG 660
DB 601 GGATCAGTGAAGCTTGTGATGTGAGTCTCAATAAAGCTCCAGGCAACAAAGCCCAAGTGAG 660
QY 661 AGCTGGACACTTCGCGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGGCAC 720
DB 661 AGCTGGACACTTCGCGTGGAGCTGGACCGGGGAGGACACAGATAAAACCCCAAGGGCAC 720
```


Db 340 GTGGCCACTTCTTCCCGGACAAACCGAGATAAGTCTGTGGATGCGCAACCGGAAAG 399
Qy 487 AATCTTTGGGAAGAGGCCAAACCCGAGCCACCGACTGCTAAATCTCGTTTCTTTGATG 546
Db 400 AATCTTTGGGAAGAGGCCAAACCCGAGCCACCGACTGCTAAATCTCGTTTCTTTGATG 459
Qy 547 CTCTCTCGGCTGTACACGAGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 606
Db 460 CTCTCTCGGCTGTACACGAGAGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 519
Qy 607 GTGAAGCTTGATGTACGTCCAATAAAGCTCCAGCGAACAAAGACCCAAAGTGAAGCTGG 666
Db 520 GTGAAGCTTGATGTACGTCCAATAAAGCTCCAGCGAACAAAGACCCAAAGTGAAGCTGG 579
Qy 667 ACACCTTCGCGTGGCAGCTGACCGGGGACGACACAGATAAACCCGAGGCGACGCCCG 726
Db 580 ACACCTTCGCGTGGCAGCTGACCGGGGACGACACAGATAAACCCGAGGCGACGCCCG 639
Qy 727 GCCCAAGACAAAGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTGAGACAGGG 786
Db 640 GCCCAAGACAAAGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTGAGACAGGG 699
Qy 787 GGAGCAGGAGAGAGCTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTC 846
Db 700 GGAGCAGGAGAGAGCTCCCTCCAGCCCAAGGACTCCAGCTTTTGTGACAAATCTTC 759
Qy 847 AAGCTGACAAAGGACAGGAAGGTGCCAGGTGACAGCCAAAGGAGCCNAGAGGGCA 906
Db 760 AAGCTGACAAAGGACAGGAAGGTGCCAGGTGACAGCCAAAGGAGCCNAGAGGGCA 819
Qy 907 GAGCATCAAGACAAAGTGGATGAGTTCTTGCGCTTATCAGGGCAGTCCGATGATGCCCT 966
Db 820 GAGCATCAAGACAAAGTGGATGAGTTCTTGCGCTTATCAGGGCAGTCCGATGATGCCCT 879
Qy 967 GCAGGGAAGGACATAGTTGACGGAAGGAAAGGAGGACAAAGAACTTGGAACTTGGCAT 1026
Db 880 GCAGGGAAGGACATAGTTGACGGAAGGAAAGGAGGACAAAGAACTTGGAACTTGGCAT 939
Qy 1027 TGCTCTGCTCCGAGGACCCAGAGGACTGGAGACTGCAAGGACGATTCACCGGACGA 1086
Db 940 TGCTCTGCTCCGAGGACCCAGAGGACTGGAGACTGCAAGGACGATTCACCGGACGA 999
Qy 1087 GCTATAGCAGAGATAATATTTCCATCATGAGTTCTTTAAACTCTGGTTTACACCTAAC 1146
Db 1000 GCTATAGCAGAGATAATATTTCCATCATGAGTTCTTTAAACTCTGGTTTACACCTAAC 1059
Qy 1147 AAAGCTGAAACAAAGGACCCAGAGACACCGGGTGTGAAAGTCAACCCACACTTCA 1206
Db 1060 AAAGCTGAAACAAAGGACCCAGAGACACCGGGTGTGAAAGTCAACCCACACTTCA 1119
Qy 1207 GCTGACCTTAAGTCAGACAAAGCCAACTTTATCCTCCAGGAGACCCAGGGCTGGCAAG 1266
Db 1120 GCTGACCTTAAGTCAGACAAAGCCAACTTTATCCTCCAGGAGACCCAGGGGCTGGCAAG 1179
Qy 1267 AATTCCAAAGGATGCAACCCATCGGGGACACACAGTCCGTCACAAACCCCTGAACCTGCG 1326
Db 1180 AATTCCAAAGGATGCAACCCATCGGGGACACACAGTCCGTCACAAACCCCTGAACCTGCG 1239
Qy 1327 AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTGCTGCTTGGGCAAACTGTTTGG 1386
Db 1240 AAGGAAGGACCAAGGAGAAATCAGGACCCACCTCTGCTGCTTGGGCAAACTGTTTGG 1299
Qy 1387 AAAAGTCAGTTAAAGGACTAGTCCCAAGTGGGAGGAGATGTTGTTGTAG 1446
Db 1300 AAAAGTCAGTTAAAGGACTAGTCCCAAGTGGGAGGAGATGTTGTTGTAG 1359
Qy 1447 TCACAGTAGAGATTATTAAGTCCAAAGGAGTGAATCAGCTTTACAAACAGTGGACCTC 1506
Db 1360 TCACAGTAGAGATTATTAAGTCCAAAGGAGTGAATCAGCTTTACAAACAGTGGACCTC 1419
Qy 1507 AACGAAGGAGATGCTGCACTTGAAACCCACAGAGCGGAAACTCAAGAGAGAAAGCAAA 1566

Db 1420 AACGAAGGAGATGCTGCACTGAAACCCACAGAACCGAAACTCAAAAGAGAGAAAGCAAA 1479
Qy 1567 CCAAGAACCTCTCTGATGGGGTTTCTCAGACAAATGTTCAGTGAAGGGGATGAGGGATC 1626
Db 1480 CCAAGAACCTCTCTGATGGGGTTTCTCAGACAAATGTTCAGTGAAGGGGATGAGGGATC 1539
Qy 1627 ACCCACTCAGAAGAAATAATGGGAAGACTCCAGCTGCCCCAAACATCAGACTCCACAGAA 1686
Db 1540 ACCCACTCAGAAGAAATAATGGGAAGACTCCAGCTGCCCCAAACATCAGACTCCACAGAA 1599
Qy 1687 AAGACTATCAGACCGCCAGAGCTGAACCAACAGGAGACCAAGAGGGTAAAGGGC 1746
Db 1600 AAGACTATCAGACCGCCAGAGCTGAACCAACAGGAGACCAAGAGGGTAAAGGGC 1659
Qy 1747 TCCTCGAAGACAAAGAGTCCAGCAGCCGAGATGAACAAAGCAGAGACAAACAGACGAA 1806
Db 1660 TCCTCGAAGACAAAGAGTCCAGCAGCCGAGATGAACAGGAGACCAACAGACGAA 1719
Qy 1807 GCCAAGAACCAAGCTCCAGAGGAGCCCAAGGAGCCCAAGGAGTCACTGCGAAT 1866
Db 1720 GCCAAGAACCAAGCTCCAGAGGAGCCCAAGGAGCCCAAGGAGTCACTGCGAAT 1779
Qy 1867 GGGGACAAAGCTCCAAAGAGACCTGAGAGCGGAGCTCCCTTGGGGGCTTTTAAA 1926
Db 1780 GGGGACAAAGCTCCAAAGAGACCTGAGAGCGGCANAGGTCCCTTGGGGGCTTTTAAA 1839
Qy 1927 GGCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGA 1986
Db 1840 GGCCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCAACAGACCCAGTATCCATCGGA 1899
Qy 1987 CCAGTTGGCAAAACCAAGTAAACAAATCAGACCGTTTCCCAAGGTTTCTCTGCGCA 2046
Db 1900 CCAGTTGGCAAAACCAAGTAAACAAATCAGACCGTTTCCCAAGGTTTCTCTGCGCA 1959
Qy 2047 AGATGTTCTCTCTTACTCTCATCTCTCTCCCAACAGCTCCATGATATATTTCTTGA 2106
Db 1960 AGATGTTCTCTCTTACTCTCATCTCTCTCCCAACAGCTCCATGATATATTTCTTGA 2019
Qy 2107 TGCCAGCAAAATGAAATTTCTGCTTAGAAAATAAGCCCGAGCTGTTGATATAGAGTGTA 2166
Db 2020 TGCCAGCAAAATGAAATTTCTGCTTAGAAAATAAGCCCGAGCTGTTGATATAGAGTGTA 2079
Qy 2167 TTATTTACGCTCTCTGCTCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGTGTTAGCAGGT 2226
Db 2080 TTATTTACGCTCTCTGCTCCAGTCTTTTCTGGCAAAATAACAGTAAAGTGTGTTAGCAGGT 2139
Qy 2227 CACCTAGTTGGTCAGAAAGTGCATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2286
Db 2140 CACCTAGTTGGTCAGAAAGTGCATGATCACCAGCAGGAGGAGGAGGAGGAGGAGGAGGAG 2199
Qy 2287 TGTGTTCCGGTTAAGTGCATGAAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCCTGT 2346
Db 2200 TGTGTTCCGGTTAAGTGCATGAAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCCTGT 2259
Qy 2347 AATCTCAGCACTTTGGGAGCCGAGGAGGAGTGCATCCTGAGGTGAGGAGTTCAAGACT 2406
Db 2360 AATCTCAGCACTTTGGGAGCCGAGGAGGAGTGCATCCTGAGGTGAGGAGTTCAAGACT 2319
Qy 2407 AGCTGGCCAACTCATGAAAACCCCGTCTCTAATAAAATACAAAATTTAGCCGAGCATG 2466
Db 2320 AGCTGGCCAACTCATGAAAACCCCGTCTCTAATAAAATACAAAATTTAGCCGAGCATG 2379
Qy 2467 GTGSCACACACCTGTAGTCCCGACTCTCGGGAGCCCAACGACGAGAACCGCTTGTACC 2526
Db 2380 GTGSCACACACCTGTAGTCCCGACTCTCGGGAGCCCAACGACGAGAACCGCTTGTACC 2439
Qy 2527 CAGGAGTGAAGTTGAGTGAAGCCGAGGAGTGCATTCATTCACCTCCACCTTGGGCGAG 2586
Db 2440 CAGGAGTGAAGTTGAGTGAAGCCGAGGAGTGCATTCACCTCCACCTTGGGCGAG 2499
Qy 2587 AGCAAGATTCTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGC 2646
Db 2500 AGCAAGATTCTATCAAAAAAAGGAGCAGTGGCAAGTAAAGTTATAGAGAGAAATGCTGC 2559


```
Qy 1456 GAGATTATAAGTCCAGGGAAGTAGAATCAGCCTTACAAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCCAGGGAAGTAGAATCAGCCTTACAAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATGCTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1575
Db 1381 GATGCTGCACCTGAACCCACAGAAAGCGAAACTCAAAAGAGAGAAAGCAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTCTCAGCAAAATGTCAGTGAAGGGGATGAGGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGCTTCTCAGCAAAATGTCAGTGAAGGGGATGAGGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAAATGGGAAGACTCCAGCTGCGCAAAACATCAGACTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAAATGGGAAGACTCCAGCTGCGCAAAACATCAGACTCCACAGAAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGCTCCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGAGGCTCCTCGAAG 1620
Qy 1756 GACAAGAGTCAGCAGCGGAGTGAACCAAGCAGAGAGCAACAGCAGGAGGCCAAGAA 1815
Db 1621 GACAAGAGTCAGCAGCGGAGTGAACCAAGCAGAGAGCAACAGCAGGAGGCCAAGAA 1680
Qy 1816 CCAGCCAGTGCAAGAGCAGGCGCACCGTGGACACGAACTCCTGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTGCAAGAGCAGGCGCACCGTGGACACGAACTCCTGCAAGATGGGGACAAG 1740
Qy 1876 CTCAAAAGAGACTGAGAGAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTGGGA 1935
Db 1741 CTCAAAAGAGACTGAGAGAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTGGGA 1800
Qy 1936 CCAAGCGGATGTTGGATGCTCAAGTCCAAACAGCAGCCAGTATCCATCGGACCAAGTGGC 1995
Db 1801 CCAAGCGGATGTTGGATGCTCAAGTCCAAACAGCAGCCAGTATCCATCGGACCAAGTGGC 1860
Qy 1996 AAACCCAGTAAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAGATGTGTT 2055
Db 1861 AAACCCAGTAAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAGATGTGTT 1920
Qy 2056 CTCCTTACTCCATCTCTCCCAACACGCTCCATGATATATATTTCTGATGCCAGCA 2115
Db 1921 CTCCTTACTCCATCTCTCCCAACACGCTCCATGATATATATTTCTGATGCCAGCA 1980
Qy 2116 AATGAATTTCTGCTAGAAATTAAGCCGAGCTGCTGATATATTCAGGTGATTTTACG 2175
Db 1981 AATGAATTTCTGCTAGAAATTAAGCCGAGCTGCTGATATATTCAGGTGATTTTACG 2040
Qy 2176 TCTCTGCTCCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTACGAGGTCACTAGTT 2235
Db 2041 TCTCTGCTCCAGTCTTTTCTGGCAATAACAGTAAAGATGGTTTACGAGGTCACTAGTT 2100
Qy 2236 GGGTCAAGAGTCGATGATCACCAAGAGGAGGAGGGAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAAGAGTCGATGATCACCAAGAGGAGGAGGGAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAGTGATGAATGCGAGTGGTGGCGGCTGGTGGCTCTCGCTGTAATCTCAGC 2355
Db 2161 GTTAAGTGATGAATGCGAGTGGTGGCGGCTGGTGGCTCTCGCTGTAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCGAGGCGAGTGGATCACCTGAGGTGAGGAGTTCAAGACTAGCTCGCC 2415
Db 2221 ACTTTGGAGGCCGAGGCGAGTGGATCACCTGAGGTGAGGAGTTCAAGACTAGCTCGCC 2280
Qy 2416 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATTAAGCAGGCGATGGTGGCAC 2475
Db 2281 AACATCATGAACCCCGTCTCTACTAAATAATACAAAATTAAGCAGGCGATGGTGGCAC 2340
Qy 2476 ACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCAGGAGTG 2535
Db 2341 ACCTGTAGTCCAGCTACTCGGAGGCCCAACGACGAGAACCGCTTGTACCCAGGAGTG 2400
```

```
Qy 2536 GAGGTTCAGTGAAGCGAAGTTGCAACATTTGCATCTCACCTTGGGCGACAGAGCAAGATT 2595
Db 2401 GAGGTTCAGTGAAGCGAAGTTGCAACATTTGCATCTCACCTTGGGCGACAGAGCAAGATT 2460
Qy 2596 CTATCAAAAAAAGGCGAGTGGCAAGTATATAGAGAGAAATCTCTCTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGCGAGTGGCAAGTATATAGAGAGAAATCTCTCTAGAGGAA 2520
Qy 2656 TTAAGCGTTCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGAGGAGACGAAAAATCC 2715
Db 2521 TTAAGCGTTCTAGTAAACGGTCTCATCTCTTAAGCTTGAAGAGGAGACGAAAAATCC 2580
Qy 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGGTTCGAAT 2775
Db 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTTGGGTGGTTCGAAT 2640
Qy 2776 TTCTTAGAACGGAATGTGTTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCGAAT 2835
Db 2641 TTCTTAGAACGGAATGTGTTGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCGAAT 2700
Qy 2836 AGGTCTCTTCTAAGTTATTGATGAGAGGAGAAAGATTAAGTCTGGGAGGGCTTAAAAATGAT 2895
Db 2701 AGGTCTCTTCTAAGTTATTGATGAGAGGAGAAAGATTAAGTCTGGGAGGGCTTAAAAATGAT 2760
Qy 2896 TTGGGAAAAACAATTTGAGGCTCAGTGACAAACGCGCAAGATTAACAATTT 2948
Db 2761 TTGGGAAAAACAATTTGAGGCTCAGTGACAAACGCGCAAGATTAACAATTT 2813
```

RESULT 4

```
US-10-240-425-385
; Sequence 385, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Bolland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; NUMBER OF SEQ ID NOS: 1598
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 385
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF041260
US-10-240-425-385
```

```
Query Match 91.7%; Score 2813; DB 18; Length 2813;
Best local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 136 CTCACCGACAGCCAGCCAGTGGSCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 195
Db 1 CTCACCGACAGCCAGCCAGTGGSCAGCAGCAGTGGAGACCCAGGACCCCTGTGCAGGAG 60
Qy 196 CAGTCCCGGTGACACAGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 255
Db 61 CAGTCCCGGTGACACAGAGGGGACTGAAGATATCTCCACAGGGGCTCAGCAGGAGCAATG 120
Qy 256 GGTAAACCAATGAGTGTTCCTCCAAAGATTTGAACCAAGAGATGAACCAAGAGCAGAG 315
```

Db 121 GGTAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGAAATGAACCAAGCAGAG 180
Qy 316 ACTTACCAAGGACACAGCGTCTGCTGAAACGGGGTCCAGTGGTGTCTGCACCCACACA 375
Db 181 ACTTACCAAGGACACAGCGTCTGCTGAAACGGGGTTCAGTGGTGTCTGCACCCACACA 240
Qy 376 GTTCAGCACTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGAGCGGATAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGGAAGTGCAGCTTGGGAATAAGTGTCAAGAGCGGATAATGTGGCCACT 300
Qy 436 TCTTCCCGGAGACAAACGGAGATAGTGTCTGTTCGGATGCAACCGGAAGAAATCTTTGGG 495
Db 301 TCTTCCCGGAGACAAACGGAGATAAGTGTCTGTTCGGATGCAACCGGAAGAAATCTTTGGG 360
Qy 496 AAAGAGCCAAACCGGAGCAACAGCTGCTAAATCTCGTTTTCTTGTGATGCTCTCTCGG 555
Db 361 AAAGAGCCAAACCGGAGCAACAGCTGCTAAATCTCGTTTTCTTGTGATGCTCTCTCGG 420
Qy 556 CCTGTACCAAGCTACCGGAGACCAAGCCCGCAGATTATCCCTTTGGATCAGTGAAGCTT 615
Db 421 CCTGTACCAAGCTACCGGAGACCAAGCCCGCAGATTATCCCTTTGGATCAGTGAAGCTT 480
Qy 616 GATGTACGCTCCAATAAGCTCCAGCGAAACAAAGACCCCAAGTGAAGCTGCGACTTCCG 675
Db 481 GATGTACGCTCCAATAAGCTCCAGCGAAACAAAGACCCCAAGTGAAGCTGCGACTTCCG 540
Qy 676 GTGCAGCTGACCGGGCAGGACACAGATAAAACCCAGGGCAGCCCGCCGCCAAGAC 735
Db 541 GTGCAGCTGACCGGGCAGGACACAGATAAAACCCAGGGCAGCCCGCCGCCAAGAC 600
Qy 736 AAGGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTGAGACAGGGGGAGCAGGA 795
Db 601 AAGGTCTCTCTGCGCCGAGGATCCACAGCTTCTCCACCTGAGACAGGGGGAGCAGGA 660
Qy 796 GGAGAAGCTCCTCCAAAGCCCAAGGACTCCAGCTTTTTTGTGACAAATCTTCAAGCTGGAC 855
Db 661 GGAGAAGCTCCTCCAAAGCCCAAGGACTCCAGCTTTTTTGTGACAAATCTTCAAGCTGGAC 720
Qy 856 AAGGGACAGGAAAGTGCAGGTGACAGCCAAACAGAGAGCCAAAGAGGCGAGAGCATCAA 915
Db 721 AAGGGACAGGAAAGTGCAGGTGACAGCCAAACAGAGAGCCAAAGAGGCGAGAGCATCAA 780
Qy 916 GACAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGGAAG 975
Db 781 GACAAGTGGATGAGTTCTGGCTTATCAGGGCAGTCCGATGATGTCCTGCGAGGGAAG 840
Qy 976 GACATAGTTGACGGCAAGGAAAGAGGACCAAGAACTTGGAACTGCGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGGCAAGGAAAGAGGACCAAGAACTTGGAACTGCGGATGCTCTGTC 900
Qy 1036 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGCGATTCAGGAGCAGCAGCTATAGCA 1095
Db 901 CCTGGGACCCAGAGGACTGGAGACTGCAAGAGCGATTCAGGAGCAGCAGCTATAGCA 960
Qy 1096 GAGATAATTAATCCATCATGAGTTCTTTAAACTCTGGTTTCACTTACCAAGCTGAA 1155
Db 961 GAGATAATTAATCCATCATGAGTTCTTTAAACTCTGGTTTCACTTACCAAGCTGAA 1020
Qy 1156 ACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCACTTCAAGCTGACCTT 1215
Db 1021 ACAAAAAGGACCCAGAGACACGGGTGCTGAAAAGTCAACCCACCACTTCAAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAAGCCAACTTTATATCCAGGAGACCAAGGGGCTGGCAAGAAATCCAAA 1275
Db 1081 AAGTCAGACAAAGCCAACTTTATATCCAGGAGACCAAGGGGCTGGCAAGAAATCCAAA 1140
Qy 1276 GGATGCAACCATCGGGCAGACACAGTCCGTGACACACCTGAACTGAACTGCGAAGAGGC 1335
Db 1141 GGATGCAACCATCGGGCAGACACAGTCCGTGACACACCTGAACTGCGAAGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260

Qy 1396 GTTAAAGGAGACTCAGTCCCCACAGGTCCGAGGAGAAATGTGGTGTGAGTCAACAGTA 1455
Db 1261 GTTAAAGGAGACTCAGTCCCCACAGGTCCGAGGAGAAATGTGGTGTGAGTCAACAGTA 1320
Qy 1456 GAGATTATAAGTCCAAAGGAAGTGAATCAGCCTTTCAAAACAGTGGACCTCAACGAGGA 1515
Db 1321 GAGATTATAAGTCCAAAGGAAGTGAATCAGCCTTTCAAAACAGTGGACCTCAACGAGGA 1380
Qy 1516 GATCTGCACCTGAAACCCACAGAAAGGAAATCTCAAAAGAGAGAAACAAACCAAGAAC 1575
Db 1381 GATCTGCACCTGAAACCCACAGAAAGGAAATCTCAAAAGAGAGAAACAAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCAACCCACTCA 1635
Db 1441 TCTCTGATGGCTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCAACCCACTCA 1500
Qy 1636 GAAGAAATAATGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAAGACTATC 1695
Db 1501 GAAGAAATAATGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAAGACTATC 1560
Qy 1696 ACACCGCAGAGCTGAAACCAACAGGAGCAACACAGAAAGGTAAAGGGCTCTCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAAACCAACAGGAGCAACACAGAAAGGTAAAGGGCTCTCTCGAAG 1620
Qy 1756 GACAAAGTTCAGCAGCGGAGATGAACAGCAGAAAGAGCAACAGCAGGAAGCCCAAGAA 1815
Db 1621 GACAAAGTTCAGCAGCGGAGATGAACAGCAGAAAGAGCAACAGCAGGAAGCCCAAGAA 1680
Qy 1816 CCAGCCAGTGCACAGAGCAGGCCACGTGTGACACGAACTCACTGCAAAATGGGGACAAG 1875
Db 1581 CCAGCCAGTGCACAGAGCAGGCCACGTGTGACACGAACTCACTGCAAAATGGGGACAAG 1740
Qy 1876 CTCAAAGAGACTGAGAAAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTTGGGA 1935
Db 1741 CTCAAAGAGACTGAGAAAGCGGACAGTCCCTTGGGGCTTCTTTAAAGGCTTGGGA 1800
Qy 1936 CCAGCGGATGTTGGATGCTCAAGTGCACAGACCCAGTATCCATCGGACCACTTGGC 1995
Db 1801 CCAGCGGATGTTGGATGCTCAAGTGCACAGACCCAGTATCCATCGGACCACTTGGC 1860
Qy 1996 AAACCAAGTAAACAAATCAGACGGTTCACACAGGTTCTCTGCCACCAAGATGCTT 2055
Db 1861 AAACCAAGTAAACAAATCAGACGGTTCACACAGGTTCTCTGCCACCAAGATGCTT 1920
Qy 2056 CTCTTACTCATCTCTCCCAACACGCTCCATATATATATTTCTTCTGATGGCCAGCA 2115
Db 1921 CTCTTACTCATCTCTCCCAACACGCTCCATATATATATTTCTTCTGATGGCCAGCA 1980
Qy 2116 AATGAAATTCGCTCAGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTACG 2175
Db 1981 AATGAAATTCGCTCAGAAATTAAGCCCGAGCTGTTGTATATTTAGAGTGTATTTACG 2040
Qy 2176 TCTCTGTCTCAGTCTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2235
Db 2041 TCTCTGTCTCAGTCTTTCTGGCAAAATAACAGTAAAGTGGTTTAGCAGGTCACTAGTT 2100
Qy 2236 GGGTCAGAAAGTTCGATGATCAACCAAGCAGAAAGGGGAAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAGAAAGTTCGATGATCAACCAAGCAGAAAGGGGAAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAAGTGAATAATGGCAGTGGTGGCCGGCTGTTGGTCTCTCGCTGTAAATCTCAGC 2355
Db 2161 GTTAAAGTGAATAATGGCAGTGGTGGCCGGCTGTTGGTCTCTCGCTGTAAATCTCAGC 2220
Qy 2356 ACTTTGGGAGCCGAGCAGGTGGATCACCTGAGTCAAGGATTCAGACTAGCTGCTGCC 2415
Db 2221 ACTTTGGGAGCCGAGCAGGTGGATCACCTGAGTCAAGGATTCAGACTAGCTGCTGCC 2280
Qy 2416 AACATCATGAAACCCCTCTCTACTAAAAATACAAAAATTAGCCAGGCGATCGTGGCACAC 2475
Db 2281 AACATCATGAAACCCCTCTCTACTAAAAATACAAAAATTAGCCAGGCGATCGTGGCACAC 2340

QY 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1395
DB 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTGGAAAAAGTCA 1260
QY 1396 GTTAAAGAGGACTCAGTCCCCACAGTGGGAGGAGAAATGTGTGTGAGTCCACAGTA 1455
DB 1261 GTTAAAGAGGACTCAGTCCCCACAGTGGGAGGAGAAATGTGTGTGAGTCCACAGTA 1320
QY 1456 GAGATTATAAGTCCAAAGAAAGTAGAATCAGCTTTTCAAAACAGTGGACCTCAACGAAGGA 1515
DB 1321 GAGATTATAAGTCCAAAGAAAGTAGAATCAGCTTTTCAAAACAGTGGACCTCAACGAAGGA 1380
QY 1516 GATCTGACCTGAACCCACAGAACCGAAAATCAAAAGAGAAAGAAAGCAACCAAGAAC 1575
DB 1381 GATCTGACCTGAACCCACAGAACCGAAAATCAAAAGAGAAAGAAAGCAACCAAGAAC 1440
QY 1576 TCTCTGATGGCGTTTCTCAGACAAATGTTCAGTGAAGGGGATGGAGGGATCAACCCACTCA 1635
DB 1441 TCTCTGATGGCGTTTCTCAGACAAATGTTCAGTGAAGGGGATGGAGGGATCAACCCACTCA 1500
QY 1636 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1695
DB 1501 GAAGAAATAAATGGGAAAGACTCCAGCTGCCAAACATCAGACTCCACAGAAAGACTATC 1560
QY 1696 ACACCGCCAGAGCTGAACCAACAGAGGACCAACAGAGGGTAAAGAGGGCTCTCGAAG 1755
DB 1561 ACACCGCCAGAGCTGAACCAACAGAGGACCAACAGAGGGTAAAGAGGGCTCTCTCGAAG 1620
QY 1756 GACAGAAGTTCAGCAGCGGAGATGAACAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1815
DB 1621 GACAGAAGTTCAGCAGCGGAGATGAACAGCAGAGAGCAACAGCAGGAAGCCAAAGAA 1680
QY 1816 CCAGCCAGTGCACAGAGCAGCCACGGTGGACACGAACTCACTCAGAAATGGGACAAAG 1875
DB 1681 CCAGCCAGTGCACAGAGCAGCCACGGTGGACACGAACTCACTCAGAAATGGGACAAAG 1740
QY 1876 CTCGAAAGAGACTTGAGAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCGCTGGGA 1935
DB 1741 CTCGAAAGAGACTTGAGAGCGGACAGTCCCTTGGGGGCTCTTTAAAGGCGCTGGGA 1800
QY 1936 CCAGAGCGGATGTGGATGCTCAAGTGCANACAGACCCAGTATCCATCGGCCAGTGGC 1995
DB 1801 CCAGAGCGGATGTGGATGCTCAAGTGCANACAGACCCAGTATCCATCGGCCAGTGGC 1860
QY 1996 AAACCAAGTAAACAAATCAGCACCGGTTCCACACAGGTTCTCTGCCACCAAGATGTGT 2055
DB 1861 AAACCAAGTAAACAAATCAGCACCGGTTCCACACAGGTTCTCTGCCACCAAGATGTGT 1920
QY 2056 CTCCTTATCCATCTCTCCCAACACAGCTCCATGTATATTTCTTGATGGCCAGCA 2115
DB 1921 CTCCTTATCCATCTCTCCCAACACAGCTCCATGTATATTTCTTGATGGCCAGCA 1980
QY 2116 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTGTATATTTAGGTGTATTTTACG 2175
DB 1981 AATGAAATTCGCTAGAAATTAAGCCGAGCTGTGTATATTTAGGTGTATTTTACG 2040
QY 2176 TCTCTGCTCCAGTCTTTTCTGGCAATTAACAGTAAAGTGGTTTACGAGGTCACTAGTT 2235
DB 2041 TCTCTGCTCCAGTCTTTTCTGGCAATTAACAGTAAAGTGGTTTACGAGGTCACTAGTT 2100
QY 2236 GGGTCAGAGAGTTCGATGATCAACCAAGCAGGAAGGGAGGAATAGAGGAATGTGTTCGG 2295
DB 2101 GGGTCAGAGAGTTCGATGATCAACCAAGCAGGAAGGGAGGAATAGAGGAATGTGTTCGG 2160
QY 2296 GTTAAGTATGAAATGGCAGTGTGTGGCGGGCGTGGTCTCGCTCTGAATCTCAGC 2355
DB 2161 GTTAAGTATGAAATGGCAGTGTGTGGCGGGCGTGGTCTCGCTCTGAATCTCAGC 2220
QY 2356 ACTTTGGAGGCCGAGGAGGAGTGAATCACTGAGTCAAGGATTCAGAGCTAGCCTGGCC 2415
DB 2221 ACTTTGGAGGCCGAGGAGGAGTGAATCACTGAGTCAAGGATTCAGAGCTAGCCTGGCC 2280

QY 2416 AACATCATGAACCCCGTCTCTACTATAAAATACAAAATTTAGCCAGGCGATGGTGCACAC 2475
DB 2281 AACATCATGAACCCCGTCTCTACTATAAAATACAAAATTTAGCCAGGCGATGGTGCACAC 2340
QY 2476 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAAACCGTTGTACCCAGGAGGTG 2535
DB 2341 ACCTGTAGTCCAGCTACTCGGGAGCCCAACGACAGAAACCGTTGTACCCAGGAGGTG 2400
QY 2536 GAGGTTCAGTGAACCCGAAAGTTGCAACATTCACCTCCACCTGGGCGACAGCAAGATT 2595
DB 2401 GAGGTTCAGTGAACCCGAAAGTTGCAACATTCACCTCCACCTGGGCGACAGCAAGATT 2460
QY 2596 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAGAAATGCTCTAGNAGAA 2655
DB 2461 CTATCAAAAAAAGGCGAGTGGCAAGTAAAGTTATAGAGAAATGCTCTAGNAGAA 2520
QY 2656 TTAAGCGTTGTAGTAAACCGCTGTCTATCTCTAAGCTTTGAAGAGGGAGACGAAATCC 2715
DB 2521 TTAAGCGTTGTAGTAAACCGCTGTCTATCTCTAAGCTTTGAAGAGGGAGACGAAATCC 2580
QY 2716 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCCGGCTGTGTGGTGGTGGTCCAAAT 2775
DB 2581 ATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCCGGCTGTGTGGTGGTGGTCCAAAT 2640
QY 2776 TTCTAGNACCGAATGTGGGTTATAGAAAAAGAAATGAATAGCTGTGTTTCAAT 2835
DB 2641 TTCTAGNACCGAATGTGGGTTATAGAAAAAGAAATGAATAGCTGTGTTTCAAT 2700
QY 2836 AGGTCTCTTGAAGTATTGATGAGAGGAGAAAGATTCAGCTGGGAGGGCTTAAATGAT 2895
DB 2701 AGGTCTCTTGAAGTATTGATGAGAGGAGAAAGATTCAGCTGGGAGGGCTTAAATGAT 2760
QY 2896 TTGGGAAACAAATTCCTTTGAGGCTCAGTGACAAACGGCAAGATTTCAACTT 2948
DB 2761 TTGGGAAACAAATTCCTTTGAGGCTCAGTGACAAACGGCAAGATTTCAACTT 2813

RESULT 6

US-10-723-860-1812
; Sequence 1812, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05982.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; PRIOR FILING DATE: 2003-11-26
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1812
; LENGTH: 2813
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1812

Query Match 91.7%; Score 2813; DB 20; Length 2813;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2813; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 136 CTCACCGACAGCAGGACCTGGGACGACCCAGCTGGAGCCAGGACCTGTGCAGAG 195
DB 1 CTCACCGACAGCAGGACCTGGGACGACCCAGCTGGAGCCAGGACCTGTGCAGAG 60
QY 196 CAGTCCCGGGTGACACGAGGGGACTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 255
DB 61 CAGTCCCGGGTGACACGAGGGGACTGAAGATCTCCACAGGGGCTCAGCAGGAGCAATG 120
QY 256 GGTAAACCAAAATGAGTGTTCCTCCCAAGAGATTGAAGCAAGAAATGAACCAAGAGCAGAG 315

Db 121 GGTAAACCAATGAGTGTCCCAAGAGTTGAAGACCAAGAGAAATGAACCAAGAGCAGAG 180
Qy 316 ACTTACCAAGACAACCGGTCTGCTGAACGGGGTTCAGTGGTGGTGTGTCACCAACACA 375
Db 181 ACTTACCAAGACAACCGGTCTGCTGAACGGGGTTCAGTGGTGGTGTGTCACCAACACA 240
Qy 376 GTTCAGCACTTAGAGAAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 435
Db 241 GTTCAGCACTTAGAGAAAGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCACT 300
Qy 436 TCTTCCCGGAGACAAACCGGAGATAGTGTCTGTTCGGGATGCCAACGGAAAGAAATCTTGGG 495
Db 301 TCTTCCCGGAGACAAACCGGAGATAGTGTCTGTTCGGGATGCCAACGGAAAGAAATCTTGGG 360
Qy 496 AAAGAGGCCAAACCGGAGGACCAAGCTGCTAAATCTCGTTTCTTCTGATGCTCTTCGG 555
Db 361 AAAGAGGCCAAACCGGAGGACCAAGCTGCTAAATCTCGTTTCTTCTGATGCTCTTCGG 420
Qy 556 CCTGTACCAAGACGTAACCGGAGACCAAGCCGAGATTCAATCCCTTGGATCAGTGAAGCTT 615
Db 421 CCTGTACCAAGACGTAACCGGAGACCAAGCCGAGATTCAATCCCTTGGATCAGTGAAGCTT 480
Qy 616 GATGTCACTCAATAAAGCTCCAGCGAAACAAAGACCAAGTGAAGCTGACACTTCCG 675
Db 481 GATGTCACTCAATAAAGCTCCAGCGAAACAAAGACCAAGTGAAGCTGACACTTCCG 540
Qy 676 GTGCAGCTGACCGGGCAGGACACAGATAAACCAGGCGACGCCCCGGCCCAAGAC 735
Db 541 GTGCAGCTGACCGGGCAGGACACAGATAAACCAGGCGACGCCCCGGCCCAAGAC 600
Qy 736 AAGTCTCTCTGCGCGCAGGAGTCCACGCTTCTCCACCTGACACAGGGGGAGCAGGA 795
Db 601 AAGTCTCTCTGCGCGCAGGAGTCCACGCTTCTCCACCTGACACAGGGGGAGCAGGA 660
Qy 796 GGAGAACTCCCTCAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 855
Db 661 GGAGAACTCCCTCAAGCCCAAGACTCCAGCTTTTGTGACAAATCTTCAAGCTGGAC 720
Qy 856 AAGGACAGGAAAGGTCCAGGTGACAGCCAAACAGGAAGCCAGAGGGCAGAGCATCAA 915
Db 721 AAGGACAGGAAAGGTCCAGGTGACAGCCAAACAGGAAGCCAGAGGGCAGAGCATCAA 780
Qy 916 GACAAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTCCAGGGAG 975
Db 781 GACAAAGTGGATGAGTTCCTGGCTTATCAGGGCAGTCCGATGATGTCCTCCAGGGAG 840
Qy 976 GACATAGTTGACGGCAAGGAAAGAGAGACAAGAACTTGGAACTGCGGGATGCTCTGTC 1035
Db 841 GACATAGTTGACGGCAAGGAAAGAGAGACAAGAACTTGGAACTGCGGGATGCTCTGTC 900
Qy 1036 CTTGGGACCCAGAGGACTGGAGACTGCNAAGGACGATTCACAGGCGAGCAGCTATAGCA 1095
Db 901 CTTGGGACCCAGAGGACTGGAGACTGCNAAGGACGATTCACAGGCGAGCAGCTATAGCA 960
Qy 1096 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTCACTTAAACAAAGCTCAA 1155
Db 961 GAGATAATAATTCATCATGAGTTCTTTTAAACTCTGGTTTCACTTAAACAAAGCTCAA 1020
Qy 1156 ACAAAAAAGGACCCAGAGAGACACGGGTGCTGAAAAGTCAACCACTTCAAGCTGACCTT 1215
Db 1021 ACAAAAAAGGACCCAGAGAGACACGGGTGCTGAAAAGTCAACCACTTCAAGCTGACCTT 1080
Qy 1216 AAGTCAGACAAGCCCACTTTATCATCCAGGAGACCCAGGGCTGGCAAGAAATTCCTCAA 1275
Db 1081 AAGTCAGACAAGCCCACTTTATCATCCAGGAGACCCAGGGCTGGCAAGAAATTCCTCAA 1140
Qy 1276 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTCCGCAAGGAGGC 1335
Db 1141 GGATGCAACCCATCGGGGACACACAGTCCGTGACAAACCCCTGAACTCCGCAAGGAGGC 1200
Qy 1336 ACCAAGGAGAAATCAGGACCCACCTCTCTGCGCTCTGGGCAAACTGTTTTGGAAAAAGTCA 1395

Db 1201 ACCAAGGAGAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGTCA 1260
Qy 1396 GTTAAAGAGGACTCAGTCCCCACACAGGTGCGGAGGAGAAATGTGGTGTGAGTCAACAGTA 1455
Db 1261 GTTAAAGAGGACTCAGTCCCCACACAGGTGCGGAGGAGAAATGTGGTGTGAGTCAACAGTA 1320
Qy 1456 GAGATTATAAGTCAAAGGAGTAGAATCAGCTTAAACAAACAGTGGACCTCAACGAAGGA 1515
Db 1321 GAGATTATAAGTCAAAGGAGTAGAATCAGCTTAAACAAACAGTGGACCTCAACGAAGGA 1380
Qy 1516 GATGCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAACCAACCAAGAAC 1575
Db 1381 GATGCTGCACCTGAACCCACAGAGCGAAACTCAAAAGAGAGAAACCAACCAAGAAC 1440
Qy 1576 TCTCTGATGGCTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1635
Db 1441 TCTCTGATGGCTTTCTCAGACAAATGTCAAGTGAAGGGGATGAGGGATCACCCACTCA 1500
Qy 1636 GAAGAAATAATGGGAAGACTCCAGCTGCGAAACATCAGACTCAACAGAAAAAGCTATC 1695
Db 1501 GAAGAAATAATGGGAAGACTCCAGCTGCGAAACATCAGACTCAACAGAAAAAGCTATC 1560
Qy 1696 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGGGCTCCCTCGAAG 1755
Db 1561 ACACCGCAGAGCTGAACCAACAGGAGCACCAAGAGGGTAAAGGGCTCCCTCGAAG 1620
Qy 1756 GACAAGAGTCAAGCAGCGGAGATCAACAGCAGAGAGCAACAGAGGAGAGCCAAAGAA 1815
Db 1621 GACAAGAGTCAAGCAGCGGAGATCAACAGCAGAGAGCAACAGAGGAGAGCCAAAGAA 1680
Qy 1816 CCAGCCAGTGCAACAGAGCAGGCGCAAGGTGACACGAACTCCTCACTGCAAGATGGGGACAAG 1875
Db 1681 CCAGCCAGTGCAACAGAGCAGGCGCAAGGTGACACGAACTCCTCACTGCAAGATGGGGACAAG 1740
Qy 1876 CTCCAAAGAGACCTGAGAGCGGCGCAGCTGCGGGCTCTTTTAAAGGCTCGGA 1935
Db 1741 CTCCAAAGAGACCTGAGAGCGGCGCAGCTGCGGGCTCTTTTAAAGGCTCGGA 1800
Qy 1936 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTGGC 1995
Db 1801 CCAAAGCGGATGTTGGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGGACCAAGTTGGC 1860
Qy 1996 AAAACCAAGTAAACAAATCAGCAAGTTCACAGGTTCTCTGCGCACCAAGATGTGTT 2055
Db 1861 AAAACCAAGTAAACAAATCAGCAAGTTCACAGGTTCTCTGCGCACCAAGATGTGTT 1920
Qy 2056 CTCTCTACTCCTCTCCCAACAGCTCCATGATATATATCTCTCTGATGGCCAGCA 2115
Db 1921 CTCTCTACTCCTCTCCCAACAGCTCCATGATATATATCTCTCTGATGGCCAGCA 1980
Qy 2116 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGATATTTAGAGTGTATTTTACG 2175
Db 1981 AATGAAATTCCTGCTAGAAATTAAGCCCGAGCTGTTGATATTTAGAGTGTATTTTACG 2040
Qy 2176 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTACAGGTCACCTAGTT 2235
Db 2041 TCTCTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTACAGGTCACCTAGTT 2100
Qy 2236 GGGTCAAGAGTCAAGTATCAACCAAGAGGAGGAGGGAATAGAGGAATGTGTTCCG 2295
Db 2101 GGGTCAAGAGTCAAGTATCAACCAAGAGGAGGAGGGAATAGAGGAATGTGTTCCG 2160
Qy 2296 GTTAAGTATGATAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCGCTCTGAATCTCAGC 2355
Db 2161 GTTAAGTATGATAAATGGCAGTGGTGGCGGGCGTGGTGGCTCTCGCTCTGAATCTCAGC 2220
Qy 2356 ACTTTGGAGGCCGAGGAGGAGTGGATCACTGAGGTCAAGGATTCAGAGCTCAAGCTTGCC 2415
Db 2221 ACTTTGGAGGCCGAGGAGGAGTGGATCACTGAGGTCAAGGATTCAGAGCTCAAGCTTGCC 2280
Qy 2416 AACATCATGAAACCCCGTCTCTCTTAAATAATACAAAAATAGCCAGGATGTTGGCACAC 2475
Db 2281 AACATCATGAAACCCCGTCTCTCTTAAATAATACAAAAATAGCCAGGATGTTGGCACAC 2340

QY 2476 ACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2535
|||||
Db 2341 ACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAGAACCGCTTGTACCCAGAGGTG 2400
|||||
QY 2536 GAGGTTGCAGTGCAGGAGGAGTGTGACCATTTGCACTCCACCTTGGGCGACAGAGCAAGATT 2595
|||||
Db 2401 GAGGTTGCAGTGCAGGAGGAGTGTGACCATTTGCACTCCACCTTGGGCGACAGAGCAAGATT 2460
|||||
QY 2596 CTATCAAAAAAAGGAGGAGTGCAGTAAGTTATAGAGAAATGCTGTAGAGGAA 2655
Db 2461 CTATCAAAAAAAGGAGGAGTGCAGTAAGTTATAGAGAAATGCTGTAGAGGAA 2520
|||||
QY 2656 TTAAGCGTTGTAGTAAACCGCTCTCATCTTCAAGCTTGAAGAGGAGGAGCAAAATCC 2715
|||||
Db 2521 TTAAGCGTTGTAGTAAACCGCTCTCATCTTCAAGCTTGAAGAGGAGGAGCAAAATCC 2580
|||||
QY 2716 ATTCTGTTAAATTTCAATCTCAAGGAGGAGAAACCGGCTGTGTTGGTGGTTGCCAAT 2775
|||||
Db 2581 ATTCTGTTAAATTTCAATCTCAAGGAGGAGAAACCGGCTGTGTTGGTGGTTGCCAAT 2640
|||||
QY 2776 TTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATPAACGTTGTTTTTCAAT 2835
Db 2641 TTCTAGAACGGAATGTGTGGGTATAGAAAAAGGAATGAATPAACGTTGTTTTTCAAT 2700
|||||
QY 2836 AGGTCCTTGTAACTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATGAT 2895
|||||
Db 2701 AGGTCCTTGTAACTTATGATGAGAGGAGAAAGATTGACTGGGAGGGCTTAAATGAT 2760
|||||
QY 2896 TTGGAAAAACAATTTGCTTTGAGGCTCAGTGACCAACGCGCAAGATTACAACTT 2948
Db 2761 TTGGAAAAACAATTTGCTTTGAGGCTCAGTGACCAACGCGCAAGATTACAACTT 2813

RESULT 7

US-10-491-213-85
; Sequence 85, Application US/10491213
; Publication No. US20050048490A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; AZIMZAI, Yalda;
; APPLICANT: BAUGHN, Mariah R.; BECHA, Shanya D.;
; APPLICANT: BOROWSKY, Mark L.; CHAWLA, Narinder K.;
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.;
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.;
; APPLICANT: GORVAD, Ann E.; GRIFFIN, Jennifer A.;
; APPLICANT: HAFALIA, April J.A.; ISON, Craig H.;
; APPLICANT: KABLE, Amy E.; KALAFUS, Daniel P.;
; APPLICANT: LEHR-MASON, Patricia M.; LU, Dzung Aina M.;
; APPLICANT: MARQUIS, Joseph P.; NGUYEN, Daniel B.;
; APPLICANT: RAMKUMAR, Jayalakmi; RICHARDSON, Thomas W.;
; APPLICANT: KAREHT, Stephanie K.; SWARNAKAR, Anita;
; APPLICANT: TANG, Y. Tom; TRAN, Uyen K.;
; APPLICANT: WARREN, Bridget A.; XU, Yuming;
; APPLICANT: YAO, Monique G.; YUE, Huibin;
; APPLICANT: YUE, Henry
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PP-1213 USN
; CURRENT APPLICATION NUMBER: US/10/491,213
; CURRENT FILING DATE: 2004-03-26
; PRIOR APPLICATION NUMBER: PCT/US02/31095
; PRIOR FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/326,389
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: US 60/327,380
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/328,186
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: US 60/329,690
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: US 60/345,384
; PRIOR FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 60/348,165
; PRIOR FILING DATE: 2001-10-26

; PRIOR APPLICATION NUMBER: US 60/350,219
; PRIOR FILING DATE: 2001-11-02
; PRIOR APPLICATION NUMBER: US 60/344,518
; PRIOR FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: US 60/345,143
; PRIOR FILING DATE: 2001-11-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PERL Program
; SEQ ID NO 85
; LENGTH: 2020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 4757551CB1
US-10-491-213-85

Query Match 62.9%; Score 1929.4; DB 21; Length 2020;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 2009; Conservative 0; Mismatches 6; Indels 66; Gaps 1;
QY 127 GGACAGAGGCTCCACCGACAGCGACTGGGAGACGCACTGGAGACCCAGACCCCT 186
Db 6 GGAGCCAGGCTCCACCGACAGCGACTGGGAGACGCACTGGAGACCCAGACCCCT 65
QY 187 GTGACGAGAGCAGCTCCCGGTGACAGGGGACTGAAGATATCTCCACAGGGGCTCAGCA 246
Db 66 GTGACGAGAGCAGCTCCCGGTGACAGGGGACTGAAGATATCTCCACAGGGGCTCAGCA 125
QY 247 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCA 306
Db 126 GGAGCAATGGGTAAACCAATGAGTGTTCCTCCCAAGAGTTGAAGACCAAGAGATGAACCA 185
QY 307 GAAGCAGAGACTTACCAAGACCAACCGCTCTGCTCTGAACGGGGTTCAGTGGTGTGTCG 366
Db 186 GAAGCAGAGACTTACCAAGACCAACCGCTCTGCTCTGAACGGGGTTCAGTGGTGTGTCG 245
QY 367 ACCCACACAGTTCAGCACTTAGAGAACTCGACTTGGGAATAAGTGTCAAGACGGATAAT 426
Db 246 ACCCACACAGTTCAGCACTTAGAGAACTCGACTTGGGAATAAGTGTCAAGACGGATAAT 305
QY 427 GTGGCCACTTCTTCCCGGACACCAACGAGAGATAGTGTGCGGATGCCAACGGAAG 486
Db 306 GTGGCCACTTCTTCCCGGACACCAACGAGAGATAGTGTGCGGATGCCAACGGAAG 365
QY 487 AATCTTGGAAAGAGGCCAAACCCGAGCACCAGCTGCTAAATCTCTGTTTTTTCTTGATG 546
Db 366 AATCTTGGAAAGAGGCCAAACCCGAGCACCAGCTGCTAAATCTCTGTTTTTTCTTGATG 425
QY 547 CTCTCTCGGCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 606
Db 426 CTCTCTCGGCTGTACCAAGGACGTACCGGAGACCAAGCCGAGATTCATCCCTTGGATCA 485
QY 607 GTGAAGCTTGATGTGCTCAGCTTCAATAAGCTCCAGGACCAAGACCAAGCTGAGAGCTGG 666
Db 486 GTGAAGCTTGATGTGCTCAGCTTCAATAAGCTCCAGGACCAAGACCAAGCTGAGAGCTGG 545
QY 667 ACCTTCCGTTGGCAGCTGGAGCCGGGACGACACAGATAAAACCCAGGGCAGCCCGCG 726
Db 546 ACCTTCCGTTGGCAGCTGGAGCCGGGACGACACAGATAAAACCCAGGGCAGCCCGCG 605
QY 727 GCCCAAGACAAAGTCTCTCTGCGGCCAGGGATCCACGCTTCTCCACCTTGAGACAGGG 786
Db 606 GCCCAAGACAAAGTCTCTCTGCGGCCAGGGATCCACGCTTCTCCACCTTGAGACAGGG 665
QY 787 GGAGCAGGAGGAGAGCTCCCTCAAGCCCAAGGACTCCAGCTTTTTTGACAAATTTCTTC 846
Db 666 GGAGCAGGAGGAGAGCTCCCTCAAGCCCAAGGACTCCAGCTTTTTTGACAAATTTCTTC 725
QY 847 AAGCTGGACAAAGGACAGGAAAGTGCAGCTGACGCCCAAGAGAGGAGGCA 906
Db 726 AAGCTGGACAAAGGACAGGAAAGTGCAGCTGACGCCCAAGAGAGGAGGCA 785

1164 QY GGACCCAGAGACACGGGTGCTGAAAGTACCCACCACTTCAGCTGACCTTTAAGTCAGA 1223
429 Db GGACCCAGAGACACGGGTGCTGAAAGTACCCACCACTTCAGCTGACCTTTAAGTCAGA 466
1224 QY CAAAGCCAACTTTTATCATCCAGAGACCCAGGGCTGCAAGAAATCCAAAGAGTGCAG 1283
467 Db GTGATGGACAAAGCTGGTCAGAGACATCCAGATCCAGGCTAGAGGACCAAGAA--AA 523
1284 QY CCCATCGGGGCACACAGATCGGTGCTGACCAACCCCTGAACTGCGAAGGAGGACCAAGGA 1343
524 Db AGCAGCTGGATAGCCCAAGGCTAGGACTGCGCTTTA----- 559
1344 QY GAAATCAGGACCCACCTCTCTGCTCTGGGCAAACTGTTTTGGAAAAAGTCAGTTAAAGA 1403
560 Db -----GAAATTTCTTAGGCTAAGGCTGCTGAAAA 590
1404 QY GGACTCAGTCCCAAGTCCAGTGGAGGAGAAATGTGTGTGAGTCCACAGTAGAGATTAT 1463
591 Db GTCAACCCACCACTTCAGCTGACCTTAAG-----TCAGACAAAGCCAACTTTACATCCCGAG 645
1464 QY AAAGTCCAAAGAGTGAATCAGCTTACAAACAGTGGACCTCAACGAGGAGGATGCTGC 1523
646 Db GAGACCCAAAGGGCTGGCA--AGAAATCCAAAGGATGCAACCCATCGGGGCACACACAGT 703
1524 QY ACCTGAAACCCACAGAAAGCGAAACTCAAAGAGAGAAAGCAACCAAGAACTCTCTGAT 1583
704 Db CGGTGACAAACCCCTGAACTCGGAAGGAGGACCAAGAGAAATCAGGACCCACCTCTC 763
1584 QY GCGCTTCTCAGACAAATGTCTAGTGAAGGGGATGGAGGGATCAACCACTCAGAGAAAT 1643
764 Db TGCCTCT----- 770
1644 QY AAATGGGAAGACTCCAGCTGCCAAATCAGACTCCACAGAAAGATATACACCGCC 1703
771 Db ---GGGCAAACTGTTTTGGAAAAAGACATCAGACTCCACAGAAAGATATACACCGCC 827
1704 QY AGAGCTCAACCAAGAGGACACACAGAGGTAAAGGGCTCTCGAAGGACAAGAA 1763
828 Db AGAGCTCAACCAAGAGGACACACAGAGGTAAAGGGCTCTCGAAGGACAAGAA 887
1764 QY GTCAGCAGCGAGATGAACAGAGCAGAGCAACAGCAGGAAGCAAGCAACAGCCCA 1823
888 Db GTCAGCAGCGAGATGAACAGAGCAGAGCAACAGCAGGAAGCAAGCAACAGCCCA 947
1824 QY GTGCACAGAGAGGCAACGGTGGACAAGAACTCACTGAGAGTGGGCAAGCTCCAAAA 1883
948 Db GTGCACAGAGAGGCAACGGTGGACAAGAACTCACTGAGAGTGGGCAAGCTCCAAAA 1007
1884 QY GAGACCTCAGAGCGGACGAGTCCCTTGGGGGCTTCTTTAAGGCTGGGACCAAGCG 1943
1008 Db GAGACCTCAGAGCGGACGAGTCCCTTGGGGGCTTCTTTAAGGCTGGGACCAAGCG 1067
1944 QY GATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGACAGTTGGCAACCCAA 2003
1068 Db GATGTTGATGCTCAAGTGCAAAACAGACCCAGTATCCATCGACAGTTGGCAATCCAA 1127
2004 QY GTAAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCAACCAAGATGTCTCTTAC 2063
1128 Db GTAAACAAATCAGCAGGTTCCCAACAGGTTCTCTGCAACCAAGATGTCTCTTAC 1187
2064 QY TCCATCTCTCCCAACACAGCTCATGTATATATCTTCTGATGGCCAGCAAGTGAAT 2123
1188 Db TCCATCTCTCCCAACACAGCTCATGTATATATCTTCTGATGGCCAGCAAGTGAAT 1247
2124 QY TCTGCTAGAAATTAAGCCGAGCTGTGTATATGAGGTGTATTTAGTCTCTCGGT 2183
1248 Db TCTGCTAGAAATTAAGCCGAGCTGTGTATATGAGGTGTATTTAGTCTCTCGGT 1307
2184 QY CCAGTCTTTTCTGCAAAATACAGTAAAGATGTTTACAGGTCACTAGTTGGGTGAGA 2243
1308 Db CCAGTCTTTTCTGCAAAATACAGTAAAGATGTTTACAGGTCACTAGTTGGGTGAGA 1367
2244 QY AGAGTCGATGATCACAAGCAGGAAAGGGAGGAATAGAGGAATGTGTTCGGGTTAAGTG 2303

1368 Db AGAGTCGATGATCACAAGCAGGAAAGGAGGAATAGAGGAATGTGTTCGGGTTAAGTG 1427
2304 QY ATGAAAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCTCTGTAAATCTCAGCACTTTGGG 2363
1428 Db ATGAAAAATGGCAGTGGTGGCCGGCGTGGTGGCTCTCGCTCTGTAAATCTCAGCACTTTGGG 1487
2364 QY AGGCGGAGCAGGTGGATCAGCTGAGGTGAGGTTCAAGACTAGCTGGCCCAACATCAT 2423
1488 Db AGGCGGAGCAGGTGGATCAGCTGAGGTGAGGTTCAAGACTAGCTGGCCCAACATCAT 1547
2424 QY GAAACCCCGTCTCTACTAAAAATACAAAAATTTAGCCAGGCATGGTGGCAGACACACCTGTAG 2483
1548 Db GAAACCCCGTCTCTACTAAAAATACAAAAATTTAGCCAGGCATGGTGGCAGACACACCTGTAG 1607
2484 QY TCCAGCTACTCTGGGAGCCCAACGACGAGAACCGCTTTGTATCCCAAGGAGTGGAGTTGC 2543
1608 Db TCCAGCTACTCTGGGAGCCCAACGACGAGAACCGCTTTGTATCCCAAGGAGTGGAGTTGC 1667
2544 QY AGTGAGCCGAAGTTGCACCATTTGCACTCCACCTGGGGCAGACAGACAGATTTCTATC-AA 2602
1668 Db AGTGAGCCGAAGTTGCACCATTTGCACTCCACCTGGGGCAGACAGACAGATTTCTATCAA 1727
2603 QY AAAAAAAGGCGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGTAGAGGAATTTAAGCG 2662
1728 Db AAAAAAAGGCGAGTGGCAAGTAAGTTATAGAGAGAAATGCTGTAGAGGAATTTAAGCG 1787
2663 QY TTGTAGTAAACCGCTGCTCATCTTAAGCTTTGAAGAGGAGAGACGAAATCCATTGTT 2722
1788 Db TTGTAGTAAACCGCTGCTCATCTTAAGCTTTGAAGAGGAGAGACGAAATCCATTGTT 1847
2723 QY TAAATTCATCTCAAGAGGAGGAGAACCCGGCTGTGTGTGGTGGTTCGCCAATTTCTTAG 2782
1848 Db TAAATTCATCTCAAGAGGAGGAGAACCCGGCTGTGTGTGGTGGTTCGCCAATTTCTTAG 1907
2783 QY AACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAATAGGGTCC 2842
1908 Db AACGGAATGTGGGGTATAGAAAAAGGAATGAATAGCGTTGTTTTCAATAGGGTCC 1967
2843 QY TTGTAAGTTATTGATGAGAGGAGGAAAGATTGACCTGGGAGGGCTTAAATGATTTGGGAA 2902
1968 Db TTGTAAGTTATTGATGAGAGGAGGAAAGATTGACCTGGGAGGGCTTAAATGATTTGGGAA 2027
2903 QY AACAAATGTTTTGAGGCTCAGTCAACCGCAAGATTAACAATTAATAAAAAA 2962
2028 Db AACAAATGTTTTGAGGCTCAGTCAACCGCAAGATTAACAATTAATAAAAAA 2087
2963 QY AAAAA 2967
2088 Db AAAAA 2092

RESULT 9

US-10-737-082-47

; Sequence 47, Application US/10737082

; Publication No. US20050130170A1

; GENERAL INFORMATION:

; APPLICANT: Bayer Healthcare LLC

; APPLICANT: Beard, Chris

; APPLICANT: Burgess, Chris

; APPLICANT: Gannon, Allison

; APPLICANT: Harvey, Jeanne

; APPLICANT: Lechner, John F.

; APPLICANT: Li, Zheng

; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

; FILE REFERENCE: 1657/2032

; CURRENT APPLICATION NUMBER: US/10/737,082

; CURRENT FILING DATE: 2003-12-16

; PRIOR APPLICATION NUMBER: US 10/737,082

; PRIOR FILING DATE: 2003-12-16

; NUMBER OF SEQ ID NOS: 300

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 47

; LENGTH: 127567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-737-082-47

Query Match 32.6%; Score 1000.6; DB 22; Length 127567;
Best Local Similarity 99.3%; Pred. No. 7.1e-269;
Matches 1016; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

Qy	1929	CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACAAACAGACCCAGTATCCATCGGACC	1988
Db	126545	CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCACAAACAGACCCAGTATCCATCGGACC	126604
Qy	1989	AGTTGGCAAAACCCAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAG	2048
Db	126605	AGTTGGCAATCCAAAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAG	126664
Qy	2049	ATGTGTTCTCTTACTCCATCTCTCCCAACACAGCTCCATGTATATATTTCTTCTGATG	2108
Db	126665	ATGTGTTCTCTTACTCCATCTCTCCCAACACAGCTCCATGTATATATTTCTTCTGATG	126724
Qy	2109	GCCAGCAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT	2168
Db	126725	GCCAGCAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT	126784
Qy	2169	ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCA	2228
Db	126785	ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCA	126844
Qy	2229	CCTAGTTGGGTTCAGAAAGTCATGATCACCAAGCAGGAAAGGAGGAATAGAGGAATG	2288
Db	126845	CCTAGTTGGGTTCAGAAAGTCATGATCACCAAGCAGGAAAGGAGGAATAGAGGAATG	126904
Qy	2289	TGTTCCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT	2348
Db	126905	TGTTCCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT	126964
Qy	2349	TCTCAGCACTTTGGGAGCCGAGGAGGTGATCACCTGAGGTCAAGAGTTCAAGACTAG	2408
Db	126965	TCTCAGCACTTTGGGAGCCGAGGAGGTGATCACCTGAGGTCAAGAGTTCAAGACTAG	127024
Qy	2409	CCTGCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT	2468
Db	127025	CCTGCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT	127084
Qy	2469	GGCACACCTGTAGTCCAGTACTCGGAGCCCAACGACGAAACCGCTGTGACCCA	2528
Db	127085	GGCACACCTGTAGTCCAGTACTCGGAGCCCAACGACGAAACCGCTGTGACCCA	127144
Qy	2529	GGAGTGGAGGTTGAGTCCAGTAAAGTTGACCAATTCACCTCCACCTGGGCGACAGAG	2588
Db	127145	GGAGTGGAGGTTGAGTCCAGTAAAGTTGACCAATTCACCTCCACCTGGGCGACAGAG	127204
Qy	2589	CAAGATTTCTATC- - -AAAAAAGGAGGAGTGGCAAGTAAAGTATAGAGGAAATGCTG	2645
Db	127205	CAAGATTTCTATCAAAAAAAGGAGGAGTGGCAAGTAAAGTATAGAGGAAATGCTG	127264
Qy	2646	CTAGAGGAATTAAGCGTTGTAGTAAACGCTGCTCATCTTAAAGCTTGAAGAGGAG	2705
Db	127265	CTAGAGGAATTAAGCGTTGTAGTAAACGCTGCTCATCTTAAAGCTTGAAGAGGAG	127324
Qy	2706	ACGAAATTCATTTGTTTAAATTCATCTCAAGGAGGAGAAACCGGGCTGTGTGAGT	2765
Db	127325	ACGAAATTCATTTGTTTAAATTCATCTCAAGGAGGAGAAACCGGGCTGTGTGAGT	127384
Qy	2766	GTTTCCCAATTTCTAGAACGGAATGTGTGGGTTATAGAAAAAGGAATGAATAGCGTTG	2825
Db	127385	GTTTCCCAATTTCTAGAACGGAATGTGTGGGTTATAGAAAAAGGAATGAATAGCGTTG	127444
Qy	2826	TTTTTCAATAGGGTCTTGTAAAGTTATGATGAGGAGGAAAGATTTGACTGGGAGGC	2885
Db	127445	TTTTTCAATAGGGTCTTGTAAAGTTATGATGAGGAGGAAAGATTTGACTGGGAGGC	127504

Qy 2886 TTAAATGATTTGGGAAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGCAAGATTACAA 2945
Db 127505 TTAAATGATTTGGGAAAAAAATTTGCTTTTGGAGCTCAGTGACAAACGCAAGATTACAA 127564
Qy 2946 CTT 2948
Db 127565 CTT 127567
RESULT 10
US-10-765-790-47
; Sequence 47, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
; APPLICANT: Bayer Healthcare LLC
; APPLICANT: Beard, Chris
; APPLICANT: Burgess, Chris
; APPLICANT: Gannon, Allison
; APPLICANT: Harvey, Jeanne
; APPLICANT: Lechner, John F.
; APPLICANT: Li, Zheng
; TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences
; FILE REFERENCE: 1657/2035
; CURRENT APPLICATION NUMBER: US/10765,790
; CURRENT FILING DATE: 2004-01-27
; PRIOR APPLICATION NUMBER: US 10/737,082
; PRIOR FILING DATE: 2003-12-16
; NUMBER OF SEQ ID NOS: 300
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 47
; LENGTH: 127567
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-765-790-47

Query Match 32.6%; Score 1000.6; DB 22; Length 127567;
Best Local Similarity 99.3%; Pred. No. 7.1e-269;
Matches 1016; Conservative 0; Mismatches 4; Indels 3; Gaps 1;
Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTGCACAAACAGACCCAGTATCCATCGGACC 1988
Db 126545 CCAGGACCAAGCGGATGTTGGATGCTCAAGTGCACAAACAGACCCAGTATCCATCGGACC 126604
Qy 1989 AGTTGGCAAAACCCAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAG 2048
Db 126605 AGTTGGCAATCCAAAGTAAACAAATCAGCACGGTTCACACAGGTTCTCTGCCACCAAG 126664
Qy 2049 ATGTGTTCTCTTACTCCATCTCTCCCAACACAGCTCCATGTATATATTTCTTCTGATG 2108
Db 126665 ATGTGTTCTCTTACTCCATCTCTCCCAACACAGCTCCATGTATATATTTCTTCTGATG 126724
Qy 2109 GCCAGCAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 2168
Db 126725 GCCAGCAATGAAATTTCTGCTAGAAATTAAGCCGAGCTGTGTATATTCAGGTGTATT 126784
Qy 2169 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCA 2228
Db 126785 ATTTACGTTCTGTGTCAGTCTTTTCTGGCAAAATAACAGTAAAGATGTTTAGCAGGTCA 126844
Qy 2229 CCTAGTTGGGTTCAGAAAGTCATGATCACCAAGCAGGAAAGGAGGAATAGAGGAATG 2288
Db 126845 CCTAGTTGGGTTCAGAAAGTCATGATCACCAAGCAGGAAAGGAGGAATAGAGGAATG 126904
Qy 2289 TGTTCCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT 2348
Db 126905 TGTTCCGGTTAAAGTATGATGAAATTCGAGTGTGCGCGGCTGTGTGTTGAGTGTATT 126964
Qy 2349 TCTCAGCACTTTGGGAGCCGAGGAGGTGATCACCTGAGGTCAAGAGTTCAAGACTAG 2408
Db 126965 TCTCAGCACTTTGGGAGCCGAGGAGGTGATCACCTGAGGTCAAGAGTTCAAGACTAG 127024
Qy 2409 CCTGCGGACATCATGAAACCCGCTCTCTACTTAAAAATACAAAAATTAAGCAGGATGGT 2468


```
Db 127025 CTGGCCAAACATCATGAAACCCCGTCTCTACTAABAAATACAAAATTAGCCAGGCATGGT 127084
Qy 2469 GGCACACACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAAACCGCTGTACCCA 2528
Db 127085 GGCACACACCTGTAGTCCAGCTACTCGGAGCCCAACGACGAAACCGCTGTACCCA 127144
Qy 2529 GGAGGTGGAGGTTCAGTGAGCCGGAAGTTGACACCTCCACCTGGGCGACAG 2588
Db 127145 GGAGGTGGAGGTTCAGTGAGCCGGAAGTTGACACCTCCACCTGGGCGACAG 127204
Qy 2589 CAAGATTCTTATC- - - AAAAAAAAAAGGCGAGTGGCAAGTAAGTTATAGAAAGAAAAATGCTG 2645
Db 127205 CAAGATTCTTATCAAAAAAAAAAGGCGAGTGGCAAGTAAGTTATAGAAAGAAAAATGCTG 127264
Qy 2646 CTAGAAAGGAATTAAGCTGTAGTAAACGCGTGTCTATCTTAAGCTTTGAAGAGGGAG 2705
Db 127265 CTAGAAAGGAATTAAGCTGTAGTAAACGCGTGTCTATCTTAAGCTTTGAAGAGGGAG 127324
Qy 2706 ACGAAATTCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGGT 2765
Db 127325 ACGAAATTCATTTGTTTAAATTCACATCTCAAGGAGGAGAACCCGGGCTGTGTGGGT 127384
Qy 2766 GGTTCGCAATTTCTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAAGCGTTG 2825
Db 127385 GGTTCGCAATTTCTTAGAACGGAATGTGTGGGTATAGAAAAGGAATGAATAAGCGTTG 127444
Qy 2826 TTTTTCAAATAGGCTTCTTGAAGTTATGATGAGAGGAAAGATTGACTGGGAGGGC 2885
Db 127445 TTTTTCAAATAGGCTTCTTGAAGTTATGATGAGAGGAAAGATTGACTGGGAGGGC 127504
Qy 2886 TTAAATGATTTGGGAAACAACTCTTTGAGGCTCAGTGACGACGCAAGATTAACA 2945
Db 127505 TTAAATGATTTGGGAAACAACTCTTTGAGGCTCAGTGACGACGCAAGATTAACA 127564
Qy 2946 CTT 2948
Db 127565 CTT 127567

RESULT 11
US-10-029-386-22674/c
; Sequence 22674, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AEMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 22674
; LENGTH: 644
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004501.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75363, EVALUE 6.00e-79
; OTHER INFORMATION: EST_HUMAN HIT: B1549901.1, EVALUE 0.00e+00
US-10-029-386-22674

Query Match 19.0%; Score 582.6; DB 16; Length 644;
Best Local Similarity 99.3%; Pred. No. 1.3e-152;
Matches 585; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy 395 TCGACTTGGGAATAAGTGTCAAGACGGATATATGTGGCCACTTTCTTCCCGGAGACAACGG 454
Db 644 TCGACTTGGGAATAAGTGTCAAGACGGATATATGTGGCCACTTTCTTCCCGGAGACAACGG 585
Qy 455 AGATAAGTGTCTTTCGGATGCCAAGCGAAGAAATCTTTGGGAAGAGGCCAAACCCGAGG 514
Db 584 AGATAAGTGTCTTTCGGATGCCAAGCGAAGAAATCTTTGGGAAGAGGCCAAACCCGAGG 525
Qy 515 CACGAGCTGTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCTGTACGAGACGCTACCG 574
Db 524 CACGAGCTGTAAATCTCGTTTTTTTCTTGATGCTCTCTCGGCTGTACGAGACGCTACCG 465
Qy 575 GAGACCAAGCCGCGAGATTTCCTTGGATCAGTGAAGCTTGTAGTCTCAGTCTCAATAAAG 634
Db 464 GAGACCAAGCCGCGAGATTTCCTTGGATCAGTGAAGCTTGTAGTCTCAGTCTCAATAAAG 405
Qy 635 CTCCAGCGAAACAAAGACCCCAAGTGTAGAGCTTGACACTTTCGCTGTGACGCTGACCGGGGC 694
Db 404 CTCCAGCGAAACAAAGACCCCAAGTGTAGAGCTTGACACTTTCGCTGTGACGCTGACCGGGGC 345
Qy 695 AGGACACAGATAAAACCCCGAGGCGACGCCCGGCCCAAGACAAGTCTCTCTGCGGCCA 754
Db 344 AGGACACAGATAAAACCCCGAGGCGACGCCCGGCCCAAGACAAGTCTCTCTGCGGCCA 285
Qy 755 GGGATCCCAACGCTTCTCCACTGTGAGACAGGGGAGAGGAGGAAAGTCTCCTCCAAGC 814
Db 284 GGGATCCCAACGCTTCTCCACTGTGAGACAGGGGAGAGGAGGAAAGTCTCCTCCAAGC 225
Qy 815 CCAAGGACTCCAGCTTTTTCACAAATCTTTCAGCTGTGACAGGACAGGAAAGTCTC 874
Db 224 CCAAGGACTCCAGCTTTTTCACAAATCTTTCAGCTGTGACAGGACAGGAAAGTCTC 165
Qy 875 CAGGTGACAGCAACAGGAAGCCAGAGGGGAGAGCATCAAGACAAGTGTGATGAGTTTC 934
Db 164 CAGGTGACAGCAACAGGAAGCCAGAGGGGAGAGCATCAAGACAAGTGTGATGAGTTTC 105
Qy 935 CTGCTTATCAGGCGAGTCCGATGATGCTCTGAGGGAAGGACATAGT 983
Db 104 CTGCTTATCAGGCGAGTCCGATGATGCTCTGAGGGAAGGTAAGTGT 56

RESULT 12
US-09-783-590-10392
; Sequence 10392, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10392
; LENGTH: 478
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (139)
; OTHER INFORMATION: n equals a,t,g, or c
```



```
; NAME/KEY: misc feature
; LOCATION: (276)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (281)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (334)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (451)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-10392

Query Match      13.0%; Score 398.6; DB 9; Length 478;
Best Local Similarity 97.0%; Pred. No. 5.1e-101;
Matches 456; Conservative 0; Mismatches 9; Indels 5; Gaps 5;

Qy 141 CCGACAGCCAGGCACTGGGCGACGACGCACTGGAGACCCAGGACCCCTGTGCGAGGACGCT 200
Db 9 CCGACAGCCAGGCACTGGGCGACGACGCACTGGAGACCCAGGACCCCTGTGCGAGGACGCT 68

Qy 201 CCGGCTGACAGAGGGGACTGAAGATCTCCCAAGGGGCTCAGCAGAGCAATGGGTAA 260
Db 69 CCGGCTGACAGAGGGGACTGAAGATCTCCCAAGGGGCTCAGCAGAGCAATGGGTAA 128

Qy 261 CCAATAGTGTCTCCCAAGAGTTGAAGACCAAGAGATGAACCCAGAGCAGACTTA 320
Db 129 CCAATNAGTNTTCCCAAGAGTTGAAGACCAAGAGATGAACCCAGAGCAGACTTA 188

Qy 321 CCAGGACAAACGGTCTGCTCTGAACGGGGTTCAGTGGTGGTGCAGCCCAACAGACTTA 380
Db 189 CCAGGACAAACGGTCTGCTCTGAACGGGGTTCAGTGGTGGTGCAGCCCAACAGACTTA 248

Qy 381 GCACCTAGAGAGTCTGACTTGGGAAT-AAAGTCAAGACGGAATATGTGGCCACTTCT 439
Db 249 GCACCTAGAGAGTCTGACTTGGGAATNAAAGTNTCAAGACGGATATGTGGCCACTTCT 308

Qy 440 CCCCGA-GACACCGGAG-TAAGTGTCTTTCGGATCCCAACGGAAGAAATCTGGGAA 497
Db 309 CCCCGAGGACACCGGAGATTAAAGTNTCAAGACGGATATGTGGCCACTTCT 368

Qy 498 AGAGGCCAAACCCGAGGACCAAGTCTGTAATCTCGTTTTTCTTTGATGCTCTCTCGGCC 557
Db 369 AGAGGCCAAACCCGAGGACCAAGTCTGTAATCTGTTTTTCTTTGATGCTCTCTCGGCC 428

Qy 558 TG-ThACAGGAGTACCGA-GACCAAGCCGAGATTCATCCCTTGGATC 605
Db 429 TGTACAGGAGTACCGGAGGNCAGCCGAGGTTCAATTCCTTGGTTC 478

RESULT 13
US-10-029-386-3914/c
; Sequence 3914, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3914
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.93
```

```
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.75
; OTHER INFORMATION: SWISSPROT HIT: O75363, EVALUE 3.00e-06
; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BE825197.1, EVALUE 0.00e+00
US-10-029-386-3914

Query Match      10.9%; Score 332.8; DB 16; Length 576;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 334; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1929 CCTGGGACCAAGCGGATGTTGGATGCTCAAGTCGCAACAGACCCAGTATCCATCGGACC 1988
Db 336 CCAGGGACCAAGCGGATGTTGGATGCTCAAGTCGCAACAGACCCAGTATCCATCGGACC 277

Qy 1989 AGTTGGCAAAACCCCAAGTAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAG 2048
Db 276 AGTTGGCAAAATCCCAAGTAACAAATCAGCAGGTTCCACACAGGTTCTCTGCCACCAAG 217

Qy 2049 ATGTGTTCTCTTACTCCATCTCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 2108
Db 216 ATGTGTTCTCTTACTCCATCTCTCTCCCAACAGCTCCATGTATATATTTCTTCTGATG 157

Qy 2109 GCCAGCAAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTTT 2168
Db 156 GCCAGCAAAATGAATTTCTGCTAGAAATTAAGCCGAGCTGTTGTATATTGAGGTGTTT 97

Qy 2169 ATTTACGTTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGTTTAGCAGGTCA 2228
Db 96 ATTTACGTTCTGCTGCTCAGTCTTTTCTGGCAAAATACAGTAAGATGTTTAGCAGGTCA 37

Qy 2229 CCTAGTTGGGTCAAGAGTCGATGATCACCACCAAGCA 2264
Db 36 CCTAGTTGGGTCAAGAGTCGATGATCACCACCAAGCA 1

RESULT 14
US-10-800-322-327/c
; Sequence 327, Application US/10800322
; Publication No. US20050053967A1
; GENERAL INFORMATION:
; APPLICANT: MEDIMOLECULAR PTY LTD
; TITLE OF INVENTION: NOVEL MARKERS AND USES THEREOF
; FILE REFERENCE: 2558321/TDO
; CURRENT APPLICATION NUMBER: US/10/800,322
; CURRENT FILING DATE: 2004-03-12
; PRIOR APPLICATION NUMBER: US 60/322228
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 338
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 327
; LENGTH: 372
; TYPE: DNA
; ORGANISM: mammalian
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: ( )..( )
; OTHER INFORMATION: "n" is an unknown nucleotide
US-10-800-322-327

Query Match      10.7%; Score 328; DB 21; Length 372;
Best Local Similarity 97.6%; Pred. No. 3e-81;
Matches 328; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 2625 AGTTATAGACAGAAATGCTCTAGAGGAATTAAGCGTTTCTAGTAACGGTCTCATC 2684
Db 356 AGTTATAGAGAGAAATGCTCTAGAGGAATTAAGCGTTTCTAGTAACGGTCTCATC 297

Qy 2685 CTCTAAGCTTGAAGAGGAGACGAAATCCATTTGTTTAAATTCACATCTCAAGAGGG 2744
Db 296 CTNTAAGCTTGAAGAGGAGACGAAATCCNTTTGTTTAAATTCACATNTCAAGAGGG 237

Qy 2745 AGAACCCGGGCTGTGTGGGTGGTTCACATTTCTTAGAACGGAATGTGGGGTATAGA 2804
```

Job time : 1794.95 secs

```
Db 236 AGAACCCGGCTGTGTGGTGGTGGTCCNATTTCTAGAACGAATGTGGGGTATAGA 177
QY 2805 AAAAGGAATGAATAGCGTGTGTTTCAATATAGGGTCTTGTAAAGTTATTGATGAGAGG 2864
Db 176 AAAAGGAATGAATAGCGTGTGTTTCAATATAGGGTCTTGTAAAGTTATTGATGAGAGG 117
QY 2865 AAAAGATTGACTGGGGGGCTTAAATGATTGGGAAACAAATTGCTTTTGAGGCTCAG 2924
Db 116 AAAAGATTGACTGGGGGGCTTAAATGATTGGGAAACAAATTGCTTTTGAGGCTCAG 57
QY 2925 TGACAAACGCAAGATTCAACTTAAAAA 2960
Db 56 TGACAAACGCAAGATTCAACTTAAAAA 21

RESULT 15
US-10-029-386-8953/c
; Sequence 8953, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 8953
; LENGTH: 506.
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004501.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.7
; OTHER INFORMATION: SWISSPROT HIT: O75363. EVALUE 3.00e-42
; OTHER INFORMATION: EST_HUMAN HIT: BG468930.1, EVALUE 1.00e-129
; OTHER INFORMATION: NT HIT: AF041260.1, EVALUE 0.00e+00
US-10-029-386-8953

Query Match 9.7%; Score 297; DB 16; Length 506;
Best Local Similarity 100.0%; Pred. No. 1.8e-72;
Matches 297; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 393 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCAATTCTTCCCGGAGACAAC 452
Db 297 AGTCGACTTGGGAATAAGTGTCAAGACGGATAATGTGGCCAATTCTTCCCGGAGACAAC 238
QY 453 GGAGATAAGTGTGTGGCGATGCCAACGGAAAGAAATCTTGGGAAAGAGGCCAAACCCGA 512
Db 237 GGAGATAAGTGTGTGGCGATGCCAACGGAAAGAAATCTTGGGAAAGAGGCCAAACCCGA 178
QY 513 GGCACACGTCTTAATCTCGTTTTTTTCTTGATGCTCTCTCGGCGCTGTACCAGGACGTAC 572
Db 177 GGCACACGTCTTAATCTCGTTTTTTTCTTGATGCTCTCTCGGCGCTGTACCAGGACGTAC 118
QY 573 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTGAGTCCCAATAA 632
Db 117 CGGAGACCAAGCCGAGATTTCATCCCTTGGATCAGTGAAGCTTGTATGTGAGTCCCAATAA 58
QY 633 AGCTCAGGGAACAAAGACCCCAAGTGAGAGCTGGACACTTCGGTGGGAGCTGGACC 689
Db 57 AGCTCAGGGAACAAAGACCCCAAGTGAGAGCTGGACACTTCGGTGGGAGCTGGACC 1
```

Search completed: July 4, 2005, 11:51:11

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:48:07 ; Search time 13508.8 Seconds

(without alignments)
11428.001 Million cell updates/sec

Title: US-08-731-499-10

Perfect score: 3186

Sequence: 1 ATGCAATCGAAAGTGACAGG.....GGAACACTACAGTGTGTAA 3186

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hhg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3183.6	99.9	3186	6	BD085734 Genes fro
2	3040.6	95.4	5632	6	AR157086 Sequence
3	3040.6	95.4	5632	9	AF041259 Homo sapi
4	3029.6	95.1	5633	6	CQ726295 Sequence
5	1561.2	49.0	128871	9	AL157838 Human DNA
6	1559.6	49.0	10365	6	BD085733 Genes fro
7	1559.6	49.0	121143	9	AF312915 Homo sapi
8	1556.4	48.9	105023	2	AC116668 Trypanoso
9	1106.2	34.7	1507	6	BD085727 Genes fro
10	1106.2	34.7	1507	6	BD085727 Genes fro
11	733	23.0	114285	10	AL844576 Mouse DNA
12	722.6	22.7	208979	2	AC023610 Mus muscu
13	716.2	22.5	184420	2	AC118373 Rattus no
14	716.2	22.5	246332	2	AC099079 Rattus no
15	444.4	13.9	469	6	BD060160 Secreter
16	315.8	9.9	2750	5	AF419155 Xenopus l
17	315.6	9.9	3512	5	BC072191 Xenopus l
18	274.4	8.6	1196	10	BC046393 Mus muscu
19	113.6	3.6	212026	2	AC136666 Rattus no

20	113.6	3.6	244851	2	AC094607 Rattus no
21	113.6	3.6	268559	2	AC112808 Rattus no
22	104.6	3.3	4934	10	BC053104 Mus muscu
23	104.6	3.3	175110	10	AC142100 Mus muscu
24	104	3.3	5982	6	CQ714295 Sequence
25	104	3.3	112797	9	AC018413 Homo sapi
26	104	3.3	178717	2	AC024396 Homo sapi
27	104	3.3	211981	9	AC009716 Homo sapi
28	102.4	3.2	6033	9	D86975 Homo sapien
29	97	3.0	8511	5	BC081149 Xenopus l
30	79.8	2.5	4265	10	BC033594 Mus muscu
31	79.8	2.5	4417	10	AK122270 Mus muscu
32	79.8	2.5	195922	10	AC140213 Mus muscu
33	77.8	2.4	243887	2	AC024898 Homo sapi
34	76.2	2.4	165094	2	CR547124 Danio rer
35	76.2	2.4	174225	5	BX005071 Zebrafish
36	76.2	2.4	192578	5	AL732455 Zebrafish
37	75.8	2.4	256281	2	AC094062 Rattus no
38	75.8	2.4	266956	2	AC133822 Rattus no
39	72	2.3	4935	6	CQ718121 Sequence
40	72	2.3	4935	9	AB002388 Homo sapi
41	72	2.3	163120	9	AC011454 Homo sapi
42	71.2	2.2	171239	2	AC150136 Gallus ga
43	71.2	2.2	190305	2	AC150054 Gallus ga
44	70	2.2	70	6	CQ876260 Sequence
45	66.8	2.1	1958	6	CQ730003 Sequence

ALIGNMENTS

RESULT 1
BD085734
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BD085734
Genes from the 20q13 amplicon and their uses.
BD085734
BD085734.1 GI:22631344
JP 2001524802-A/10.
synthetic construct
other sequences; artificial sequences.
1 (bases 1 to 3186)
Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and
Rommsen,J.
Patent: JP 2001524802-A 10 04-DEC-2001;
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS
PN
PD
PF
PR
PI
PI
PC
CC
Key
FT
FT

3186 bp
DNA
linear
PAT 27-AUG-2002

17-JAN-1997 US 08/785532
JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
JOHANNA ROMMENS
C12Q1/68, A61K48/00
Description of Artificial Sequence: ZABC1 Open Reading Frame FH
Location/Qualifiers
1..3186
Location/Qualifiers
1..3186
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

Query Match 99.9%; Score 3183.6; DB 6; Length 3186;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCNATCGAAGTGACAGGAAACATGCCACTCAATCCCTCTTAATGTACATGATGGG 60

Db 1 ATGCAATCGAAAGTGACAGGAAAATGTCACAACTCAATCCCTCTTAATGTACATGATGGG 60
Qy
Db 61 CCAGAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTCTCAATG 120
Db 61 CCAGAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTCTCAATG 120
Qy 121 AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG 180
Db 121 AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG 180
Qy 181 GGGTATATGCTTGGATTTGATTTGAGTTCGAGCAGACCTTCAACATTCAGAGACCTT 240
Db 181 GGGTATATGCTTGGATTTGATTTGAGTTCGAGCAGACCTTCAACATTCAGAGACCTT 240
Qy 241 AATAAATGCTTTAATGCAACACGGGCTTACCCTCTGTGAACGAGCATTTCTTCGGGTT 300
Db 241 AATAAATGCTTTAATGCAACACGGGCTTACCCTCTGTGAACGAGCATTTCTTCGGGTT 300
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGCAACAGAACCTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGGCAACAGAACCTCCCAAGGAA 360
Qy 361 AAGAATTCGAAGGAAAAATGAAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGCT 420
Db 361 AAGAATTCGAAGGAAAAATGAAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGCT 420
Qy 421 TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAACATG 480
Db 421 TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAACATG 480
Qy 481 TGGGGAAGAGMTTSSRSAGCTTGGTTTCTTAAATAATCAATGCGGACACATAATGGC 540
Db 481 TGGGGAAGAGMTTCAAGAGGCTTGGTTTCTTAAATAATCAATGCGGACACATAATGGC 540
Qy 541 AAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 541 AAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Qy 601 GAGTCTGTCAGGTGCAACGGGCGAGAGCATCTCTCTTCAAAATCTGCATGGTT 660
Db 601 GAGTCTGTCAGGTGCAACGGGCGAGAGCATCTCTCTTCAAAATCTGCATGGTT 660
Qy 661 TGTGCTCTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAGCGCAAGGTGCAACACAAA 720
Db 661 TGTGCTCTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCAGCGCAAGGTGCAACACAAA 720
Qy 721 AAAACTGCTTTTGGTACACAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC 780
Db 721 AAAACTGCTTTTGGTACACAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC 780
Qy 781 TCGAGGGAGGACTTCTGCAAGTGTTCACCTTGAGACCAAAATCTCACCTGMAACGGGG 840
Db 781 TCGAGGGAGGACTTCTGCAAGTGTTCACCTTGAGACCAAAATCTCACCTGMAACGGGG 840
Qy 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCCACACCTTCCAGGCTTGGCAG 900
Db 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTCCACACCTTCCAGGCTTGGCAG 900
Qy 901 CTGGCTACCAAGGAAAAAGTTGCCATTTTCCAAAGAAAGTGAAGGAATCGGGGCAAGAGGG 960
Db 901 CTGGCTACCAAGGAAAAAGTTGCCATTTTCCAAAGAAAGTGAAGGAATCGGGGCAAGAGGG 960
Qy 961 AGCACCGACAAGAGGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGGCAAT 1020
Db 961 AGCACCGACAAGAGGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGGGCAAT 1020
Qy 1021 TGTGAGGCTCTCCCAAGAGAAAGAGAAAGTGCACCAACTCCACGGGCAAGCGCCCTCC 1080
Db 1021 TGTGAGGCTCTCCCAAGAGAAAGAGAAAGTGCACCAACTCCACGGGCAAGCGCCCTCC 1080
Qy 1081 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGGAGAAAGCCCACTCACTGCTCGAGTGC 1140

Db 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGCAAGAGAGGCCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAGAGCTTTCAAACTTACACAGCTGTCTTGCATCTCCAGGCTCCCAAGAAAGAC 1200
Db 1141 GGCAAGAGCTTTCAAACTTACACAGCTGTCTTGCATCTCCAGGCTCCCAAGAAAGAC 1200
Qy 1201 CGGAGGCGCGCGGAGTCCGCCACCATGTCTGTGAACGGAGCGACCGGGGACGCTGT 1260
Db 1201 CGGAGGCGCGCGGAGTCCGCCACCATGTCTGTGAACGGAGCGACCGGGGACGCTGT 1260
Qy 1261 TCTCTGACCTCGCGGCTCTCGATGAAATCGAGCCGCTGGATCGAGGGGAAAGGTGT 1320
Db 1261 TCTCTGACCTCGCGGCTCTCGATGAAATCGAGCCGCTGGATCGAGGGGAAAGGTGT 1320
Qy 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAGAGGAATCCATCTGGATAAATAATGATGAT 1380
Db 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAGAGGAATCCATCTGGATAAATAATGATGAT 1380
Qy 1381 GGAGGAAAAATAAACATCTTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1440
Db 1381 GGAGGAAAAATAAACATCTTTACATCTTCAAGAGAGTGTAGTTATTTGTGGAAGTTTTTC 1440
Qy 1441 CGTTCAAATTTATTAACCTCAATATTTCAATCTCAGAACGATACAGGTGAAACCAATACAAA 1500
Db 1441 CGTTCAAATTTATTAACCTCAATATTTCAATCTCAGAACGATACAGGTGAAACCAATACAAA 1500
Qy 1501 TGTGAATTTTGTGAATATGCTGAGCCGAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1501 TGTGAATTTTGTGAATATGCTGAGCCGAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Qy 1561 CATCAAGGAAAAACCAACCGATTTGCTGCTGAGGTCAAGAACGATGGTAAATAATCAG 1620
Db 1561 CATCAAGGAAAAACCAACCGATTTGCTGCTGAGGTCAAGAACGATGGTAAATAATCAG 1620
Qy 1621 GACACTGAAGATGCACTATTAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1680
Db 1621 GACACTGAAGATGCACTATTAACCGCTGACAGTGCAGAACCAAAATTTGAAAGATTT 1680
Qy 1681 TTTGATGCTGCAAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAAGAGATGCT 1740
Db 1681 TTTGATGCTGCAAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAAGAGATGCT 1740
Qy 1741 TCTGTTTTTCAAGATTTTGGGCGAGCTGCTCTCTCACAGCAGACACAAAGATATCTCAG 1800
Db 1741 TCTGTTTTTCAAGATTTTGGGCGAGCTGCTCTCTCACAGCAGACACAAAGATATCTCAG 1800
Qy 1801 GATTTCATATAAATGACGCTGATGACGCTGTGATAAGTGAATAAATAACCTACCCCT 1860
Db 1801 GATTTCATATAAATGACGCTGATGACGCTGTGATAAGTGAATAAATAACCTACCCCT 1860
Qy 1861 GCTTACCTGGACCTTTTAAATAAGAGATCAGCAGTGTGAACTCAGGCAAAATAACCTCATC 1920
Db 1861 GCTTACCTGGACCTTTTAAATAAGAGATCAGCAGTGTGAACTCAGGCAAAATAACCTCATC 1920
Qy 1921 TGTAGAACCAAGGCGGATTTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 1980
Db 1921 TGTAGAACCAAGGCGGATTTTACTCTCTCCGATGGCAGTACACCCATAACCTTGAA 1980
Qy 1981 GTTAGCCCCCAAGAGAGCAACCGAGACCGAGCTGATGAGTAAAGTGAATAAATAACCTTG 2040
Db 1981 GTTAGCCCCCAAGAGAGCAACCGAGACCGAGCTGATGAGTAAAGTGAATAAATAACCTTG 2040
Qy 2041 GATTGTACGAAAAACCTTTTAAATTTATTCGTTGGGGCTTTTCAAAATTTGCCCGCAAT 2100
Db 2041 GATTGTACGAAAAACCTTTTAAATTTATTCGTTGGGGCTTTTCAAAATTTGCCCGCAAT 2100
Qy 2101 TCTTTGAGTAAAAAGTTTGAATTCGAAGTATACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2101 TCTTTGAGTAAAAAGTTTGAATTCGAAGTATACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Qy 2161 TATCCAGAGGTTTAAATGATGACACGAGACTGGAGCATTAATAAATCTTGACGCTTCAT 2220
Db 2161 TATCCAGAGGTTTAAATGATGACACGAGACTGGAGCATTAATAAATCTTGACGCTTCAT 2220

```
QY 2221 AAAAAGTCTGAAACAGTCCCTGCTTGAAGTCAAGTACCGATGCCCGCAGCGTTG 2280
DB 2221 AAAAAGTCTGAAACAGTCCCTGCTTGAAGTCAAGTACCGATGCCCGCAGCGTTG 2280
QY 2281 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
DB 2281 CTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
QY 2341 CCGGCGAGTCAAAATCCCTGCTGATCTCGGAAGGGGAAAGAGCCCTCTCGGGCAGGC 2400
DB 2341 CCGGCGAGTCAAAATCCCTGCTGATCTCGGAAGGGGAAAGAGCCCTCTCGGGCAGGC 2400
QY 2401 AAGGCCCTCTGACTTCAGGATGAGCTAGCACTTTAGCCCAAGTAAGCTGAAGTCC 2460
DB 2401 AAGGCCCTCTGACTTCAGGATGAGCTAGCACTTTAGCCCAAGTAAGCTGAAGTCC 2460
QY 2461 CACAGACACAGCAGAAATGTGGGGTCCAAGGGGCGGCACAGCAACAGCAATCTGAG 2520
DB 2461 CACAGACACAGCAGAAATGTGGGGTCCAAGGGGCGGCACAGCAACAGCAATCTGAG 2520
QY 2521 ATGTTTCTAAACCAAGTGTTCCTCTGACCGGATAGACAAAAGACCCGAGACAAA 2580
DB 2521 ATGTTTCTAAACCAAGTGTTCCTCTGACCGGATAGACAAAAGACCCGAGACAAA 2580
QY 2581 TTGAAACCTCTTCAGTAGCTCTCTCAGCGCCACCTCGGGAGAGACTATTTTC 2640
DB 2581 TTGAAACCTCTTCAGTAGCTCTCTCAGCGCCACCTCGGGAGAGACTATTTTC 2640
QY 2641 TCCATCGACTACCCCGCAAGAACAGACCGCTGGGCACCTCGGGAGAGACTATTTTC 2700
DB 2641 TCCATCGACTACCCCGCAAGAACAGACCGCTGGGCACCTCGGGAGAGACTATTTTC 2700
QY 2701 TGTAAATCGGAGTCCAGCAATCTGACAGCAAAATTTGTGAGCCCTTCCAAAAGACTG 2760
DB 2701 TGTAAATCGGAGTCCAGCAATCTGACAGCAAAATTTGTGAGCCCTTCCAAAAGACTG 2760
QY 2761 AAGTCCAGCTGTGCTGCTTGAAGTCAAGCCCGGGGCAATACAGAGAGGCTAT 2820
DB 2761 AAGTCCAGCTGTGCTGCTTGAAGTCAAGCCCGGGGCAATACAGAGAGGCTAT 2820
QY 2821 GACCTTCCCAAGTACCATATGCTCAGAGGCAATCAGTCACTGTTACCGCAGAGCTGTG 2880
DB 2821 GACCTTCCCAAGTACCATATGCTCAGAGGCAATCAGTCACTGTTACCGCAGAGCTGTG 2880
QY 2881 TATCCGTCGAGGCGCTGCCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTCT 2940
DB 2881 TATCCGTCGAGGCGCTGCCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTCT 2940
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCCGGGCGCACTTACACTGTGTG 3000
DB 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCCGGGCGCACTTACACTGTGTG 3000
QY 3001 CCTGCTGTGACTGCTGAGTCCAGCTCAGCTTGAAGGCTTGTGAGTGTGAGTGTGCTTA 3060
DB 3001 CCTGCTGTGACTGCTGAGTCCAGCTCAGCTTGAAGGCTTGTGAGTGTGAGTGTGCTTA 3060
QY 3061 CTCCCAAGAAATTAATTTTACTTCTATCTTTCAGAGCGAAATGGTGAAGTCTACTGAA 3120
DB 3061 CTCCCAAGAAATTAATTTTACTTCTATCTTTCAGAGCGAAATGGTGAAGTCTACTGAA 3120
QY 3121 ATAAGCTGTGATGTGACTGATCAATAAATCATATGAGGAATCTGCAAGGAAACACTACGTT 3180
DB 3121 ATAAGCTGTGATGTGACTGATCAATAAATCATATGAGGAATCTGCAAGGAAACACTACGTT 3180
QY 3181 GTGTAA 3186
DB 3181 GTGTAA 3186
```

RESULT 2
AR157086
LOCUS AR157086 5632 bp DNA linear PAT 08-AUG-2001

```
DEFINITION Sequence 3 from patent US 6242590.
ACCESSION AR157086
VERSION AR157086.1 GI:15125790
KEYWORDS
SOURCE
ORGANISM
REFERENCE
  1 (bases 1 to 5632)
  Cowbert, L.M.
  Antisense modulation of zinc finger protein-217 expression
  Patent: US 6242590-A 3 05-JUN-2001;
  Location/Qualifiers
    source
      1..5632
      /organism="unknown"
      /mol_type="unassigned DNA"

ORIGIN
  Query Match          95.4%; Score 3040.6; DB 6; Length 5632;
  Best Local Similarity 95.8%; Pred. No. 0;
  Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

QY 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 60
DB 272 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 331
QY 61 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
DB 332 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 391
QY 121 AAAGGGACCGCTGTTTCCATTCGAGCTACACAAGAAAGAAATGTCCATCCAAATCGAG 180
DB 392 AAAGGGACCGCTGTTTCCATTCGAGCTACACAAGAAAGAAATGTCCATCCAAATCGAG 451
QY 181 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAGACCTT 240
DB 452 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAGACCTT 511
QY 241 AATAAACATGTTCTTAATGCAACACCGGCTACCTCTGTGAACCAAGAGTTCTTCGGGTT 300
DB 512 AATAAACATGTTCTTAATGCAACACCGGCTACCTCTGTGAACCAAGAGTTCTTCGGGTT 571
QY 301 GAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCGAAACAGAACCTCCCAAGGAA 360
DB 572 GAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCGAAACAGAACCTCCCAAGGAA 631
QY 361 AAGAATTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 420
DB 632 AAGAATTGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 691
QY 421 TTTGATGTTGAGATCCCATGAGAACACACAAGATTTCTTTCACTTACGGGTGTAAATG 480
DB 692 TTTGATGTTGAGATCCCATGAGAACACACAAGATTTCTTTCACTTACGGGTGTAAATG 751
QY 481 TGCGGAAGAGAGTTTSSAGCCCTTGTCTTAAATTCACATGCGGACACATTAATGGC 540
DB 752 TGCGGAAGAGAGTTTSSAGCCCTTGTCTTAAATTCACATGCGGACACATTAATGGC 811
QY 541 AAATCGGGGGCCAGAGCAAACTGTCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 600
DB 812 AAATCGGGGGCCAGAGCAAACTGTCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC 871
QY 601 GAGTCGTGTCAGGTGCAAGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 660
DB 872 GAGTCGTGTCAGGTGCAAGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 931
QY 661 TGTGGCTCTCTTATTTCCAAATTAAGAAAGTCTTAATTTAGAGCCCGCAAGTGCACACCAA 720
DB 932 TGTGGCTCTCTTATTTCCAAATTAAGAAAGTCTTAATTTAGAGCCCGCAAGTGCACACCAA 991
QY 721 AAAAAGTGTCTTCCGTTACCAGCAGCGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTC 780
DB 992 AAAAAGTGTCTTCCGTTACCAGCAGCGCGCAGACAGACTCTCCACAGGAGGAAATGCCGTC 1051
```

QY 781 TCAGGAGGAGCTTCCTCAGTGTGTTCAACTTGAGACCAAAATCTCAACCTGAAACGGG 840
Db TCAGGAGGAGCTTCCTCAGTGTGTTCAACTTGAGACCAAAATCTCAACCTGAAACGGG 1111
QY 841 AAGAAGCTGTCAAGTGCATCCCTCAGCTCGATCGTTCACACCTTCAGGCTTGGCAG 900
Db AAGAAGCTGTCAAGTGCATCCCTCAGCTCGATCGTTCACACCTTCAGGCTTGGCAG 1171
QY 901 CTGGCTACCAAGGAAAGTGGCAATTTGCCAATGCAAGAGTGAAGGATCGGGCAAGAGG 960
Db CTGGCTACCAAGGAAAGTGGCAATTTGCCAATGCAAGAGTGAAGGATCGGGCAAGAGG 1231
QY 961 AGCACCGACACGAGTTCGAGTTCGAGAGGAGCTTTGGAGAAAACAAATAAGGGCAGT 1020
Db AGCACCGACACGAGTTCGAGTTCGAGAGGAGCTTTGGAGAAAACAAATAAGGGCAGT 1291
QY 1021 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAAACCTCCAGCGGCGAAGCGCCCTCC 1080
Db TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCAAAACCTCCAGCGGCGAAGCGCCCTCC 1351
QY 1081 GTGAGCGGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db GTGAGCGGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCCACTCACTGCTCCGAGTGC 1411
QY 1141 GGCAAAGCTTTCAGAACCTACACAGCTGGTCTTTGCACTCCAGGGTCCACAAGAGGAC 1200
Db GGCAAAGCTTTCAGAACCTACACAGCTGGTCTTTGCACTCCAGGGTCCACAAGAGGAC 1471
QY 1201 CGGAGGCGGCGGAGTGCGCCACCATGTCTGTGGACGGGAGCGACCGGGGACGTGT 1260
Db CGGAGGCGGCGGAGTGCGCCACCATGTCTGTGGACGGGAGCGACCGGGGACGTGT 1531
QY 1261 TCTCTGACCTCGCGCCCTCTGGATGAAATGAGCGCTGGATCGAGGGAGGTGGT 1320
Db TCTCTGACCTCGCGCCCTCTGGATGAAATGAGCGCTGGATCGAGGGAGGTGGT 1591
QY 1321 TCTGAAGACGGATCTGAGGATGGGCTTCCGAAGGAATCCATCTGGATAAAATATGATGAT 1380
Db TCTGAAGACGGATCTGAGGATGGGCTTCCGAAGGAATCCATCTGGATAAAATATGATGAT 1651
QY 1381 GGAGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAAGTTTTTC 1440
Db GGAGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAAGTTTTTC 1711
QY 1441 GCTTCAAAATTTACCTCAATATTCATCTCAGACGCATACAGGTGAAAACCCATACAAA 1500
Db GCTTCAAAATTTACCTCAATATTCATCTCAGACGCATACAGGTGAAAACCCATACAAA 1771
QY 1501 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA 1560
Db TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA 1831
QY 1561 CATCAAGGAAAAACAAACCGATGTTGCTGTGAAGTCAAGACGATGGTAAATCAG 1620
Db CATCAAGGAAAAACAAACCGATGTTGCTGTGAAGTCAAGACGATGGTAAATCAG 1891
QY 1621 GACACTGAAGTGCATTAACCCCTGACAGTGCAGCAACCAAAATTTGAAAAGATTT 1680
Db GACACTGAAGTGCATTAACCCCTGACAGTGCAGCAACCAAAATTTGAAAAGATTT 1951
QY 1681 TTTGATGTTGCCAAAGATGTTACAGCGAGTCCACCTGCAAGACGCTTAAAGGAGTGCCT 1740
Db TTTGATGTTGCCAAAGATGTTACAGCGAGTCCACCTGCAAGACGCTTAAAGGAGTGCCT 2011
QY 1741 TCTGTTTTTCAAGATGTTTGGGACGCTGTCCTCTCAACAGCAACAAAGATACTCAG 1800
Db TCTGTTTTTCAAGATGTTTGGGACGCTGTCCTCTCAACAGCAACAAAGATACTCAG 2071
QY 1801 GATTTCCATAAATGCAAGCTGATCAGAGTGTGATTAAGTAATAAAACCCCTACCCCT 1860
Db GATTTCCATAAATGCAAGCTGATCAGAGTGTGATTAAGTAATAAAACCCCTACCCCT 2131
QY 1861 GCTTACCTGGACCTGTAAAAAAGAGATCAGCAGTTGAAATCTCAGGCAAAATAACCTCATC 1920

Db 2132 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAATCTCAGGCAAAATAACCTCATC 2191
QY 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAGTACACCCATAAACCCTTCAA 1980
Db TGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAGTACACCCATAAACCCTTCAA 2251
QY 1981 GTTAGCCCCCAAGAGAAAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTG 2040
Db GTTAGCCCCCAAGAGAAAGCAAAACGAGACCGCAGCTGACTGCAGATACAGGCCCAAGTGTG 2311
QY 2041 GATTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATGTCGCGCAATTT 2100
Db GATTGTCAAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATGTCGCGCAATTT 2371
QY 2101 TCTTTGAGTAAAGTGTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db TCTTTGAGTAAAGTGTGATTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2431
QY 2161 TATCAGAAAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAATCTCGACGTTTCA 2220
Db TATCAGAAAGTTTAAATGATGCACAGAGACTGGAGCATAAATACAAATCTCGACGTTTCA 2491
QY 2221 AAAAACTGTGAAACAAAGTCTTCTGTAGAAAGTCCAGTACCCGATGCCGCGAGCTTG 2280
Db AAAAACTGTGAAACAAAGTCTTCTGTAGAAAGTCCAGTACCCGATGCCGCGAGCTTG 2551
QY 2281 CTGGAAAAAGATGTGCCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
Db CTGGAAAAAGATGTGCCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2611
QY 2341 CCGCGCAGTCCAAATCCCTGCCATCTCGAAGGGAGCAGAGCCCTCTCTGGGCGAGGC 2400
Db CCGCGCAGTCCAAATCCCTGCCATCTCGAAGGGAGCAGAGCCCTCTCTGGGCGAGGC 2671
QY 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCAGTCTTTAGCCCCCAAGTAACTGAAGTCC 2460
Db AAGGCCCTCTGACTTCAGGGATAGACTCTAGCAGTCTTTAGCCCCCAAGTAACTGAAGTCC 2731
QY 2461 CACAGACACAGCAGATGTGGGGTCCAGGGGCGCCACCGAGCACACCAATCTCAG 2520
Db CACAGACACAGCAGATGTGGGGTCCAGGGGCGCCACCGAGCACACCAATCTCAG 2791
QY 2521 ATGTTTCTAAAAACAGTGTTCCTCTGCAACGGATAGACAAAAGAACCCGAGACAAA 2580
Db ATGTTTCTAAAAACAGTGTTCCTCTGCAACGGATAGACAAAAGAACCCGAGACAAA 2851
QY 2581 TTGAAAACCTTTCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGTT 2640
Db TTGAAAACCTTTCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGTT 2911
QY 2641 TCCATCGACTACCCCGCAAGAACGACAGCCCTGGGCACTCTCCGGGAAGAGACTATTTC 2700
Db TCCATCGACTACCCCGCAAGAACGACAGCCCTGGGCACTCTCCGGGAAGAGACTATTTC 2971
QY 2701 TGTAACTCGAGTGCAGCAATACTGACAGAAATTTGTTGAGCCCTTCCAAAAAGACTG 2760
Db TGTAACTCGAGTGCAGCAATACTGACAGAAATTTGTTGAGCCCTTCCAAAAAGACTG 3031
QY 2761 AAGTCCAGCTGTTGCTTGAAGTTCAGCCCGGGGCAATTAACAGAAAGAGCTAT 2820
Db AAGTCCAGCTGTTGCTTGAAGTTCAGCCCGGGGCAATTAACAGAAAGAGCTAT 3091
QY 2821 GACCTTCCAGTAGTACCATATGTCAGAGCATACATCACTGTTTACCCGAGACTGTG 2880
Db GACCTTCCAGTAGTACCATATGTCAGAGCATACATCACTGTTTACCCGAGACTGTG 3151
QY 2881 TATCGTCCGAGGCGCTGCCCTCCCAAAACCAAGTTTCTCTGAGCTCCAGCGAGTCTC 2940
Db TATCGTCCGAGGCGCTGCCCTCCCAAAACCAAGTTTCTCTGAGCTCCAGCGAGTCTC 3211
QY 2941 CCAATGTGCTGATGTTTCAAGGCCCTTATGTTGCTCCGGGCACTTTTACATTTGTTG 3000

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAAGGAA 360
Db |||||
572 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGACAGAACCTCCCAAGGAA 631
Qy |||||
361 AAGAAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGCCT 420
Db |||||
632 AAGAAATTGCAAGGAAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGCCT 691
Qy |||||
421 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Db |||||
692 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 751
Qy |||||
481 TCGGGAAGAAAGTTTSSRSGACCTTTGTTTCTTAAAAATCACATCGGCACACATAATGGC 540
Db |||||
752 TCGGGAAGAAATTCAGGAGCCTTGGTTTCTTAAAAATCACATCGGCACACATAATGGC 811
Qy |||||
541 AAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGATGATGCCAGCAACGATCAAC 600
Db |||||
812 AAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGATGATGCCAGCAACGATCAAC 871
Qy |||||
601 GAGGTGCTGCCAGGTGCACGCGCCGAGAGCATCTCTCTCTTACAAATCTGCATGGTT 660
Db |||||
872 GAGGTGCTGCCAGGTGCACGCGCCGAGAGCATCTCTCTCTTACAAATCTGCATGGTT 931
Qy |||||
661 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA 720
Db |||||
932 TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA 991
Qy |||||
721 AAAACTGCTTTGGTATCCAGCAGCGCGCAGACAGACTCTCCACAAAGGAGGAATGCCGTCC 780
Db |||||
992 AAAACTGCTTTGGTATCCAGCAGCGCGCAGACAGACTCTCCACAAAGGAGGAATGCCGTCC 1051
Qy |||||
781 TCGAGGAGGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACCTTGAACGGGG 840
Db |||||
1052 TCGAGGAGGACTTCTCAGTGTGTTCAACTTGAGACCAAAATCTCACCTTGAACGGGG 1111
Qy |||||
841 AAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG 900
Db |||||
1112 AAGAGCCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG 1171
Qy |||||
901 CTGGCTACAAAGGAAAAGTTGCCATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAGGG 960
Db |||||
1172 CTGGCTACAAAGGAAAAGTTGCCATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAGGG 1231
Qy |||||
961 AGCACGACACAGCAGATTCGAGTTCCGAGAGGAGCTTGAGAACAAATAAGGGCAGT 1020
Db |||||
1232 AGCACGACACAGCAGATTCGAGTTCCGAGAGGAGCTTGAGAACAAATAAGGGCAGT 1291
Qy |||||
1021 TGTGAGGCTCTCGCAAGAGAAAAGAGAGTGCAAACACTCCACGCGGAAGCGCCCTCC 1080
Db |||||
1292 TGTGAGGCTCTCGCAAGAGAAAAGAGAGTGCAAACACTCCACGCGGAAGCGCCCTCC 1351
Qy |||||
1081 GTGAGCGGGATCCCAAGTTACCAAGTAGCAAGGAGAACCCACTCACTGTCTCGAGTGC 1140
Db |||||
1352 GTGAGCGGGATCCCAAGTTACCAAGTAGCAAGGAGAACCCACTCACTGTCTCGAGTGC 1411
Qy |||||
1141 GGCAGAGCTTTCAGAACCTACCAACAGCTGCTTTCGACTCCAGGGTCCACAAAGAGAC 1200
Db |||||
1412 GGCAGAGCTTTCAGAACCTACCAACAGCTGCTTTCGACTCCAGGGTCCAAAGAGGAC 1471
Qy |||||
1201 CGGAGGCGGGCGGGAGTCCGCCACCATGTCTGTGGAGGGAGGAGCGGGGACGTGT 1260
Db |||||
1472 CGGAGGCGGGCGGGAGTCCGCCACCATGTCTGTGGAGGGAGGAGCGGGGACGTGT 1531
Qy |||||
1261 TCTCTGACCTCGCGCCCTCTGATGAAATGAGCGCGTGGATCGAGGGGAAGGTGT 1320
Db |||||
1532 TCTCTGACCTCGCGCCCTCTGATGAAATGAGCGCGTGGATCGAGGGGAAGGTGT 1591
Qy |||||
1321 TCTGAGACGATCTGAGATCGGCTTCCGAGGAATCCATCTCGATAAAATGATGAT 1380
Db |||||
1592 TCTGAGACGATCTGAGATCGGCTTCCGAGGAATCCATCTCGATAAAATGATGAT 1651
Qy |||||
1381 GGAGGAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGTTTTTC 1440

Db |||||
1652 GGAGGAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGTTTTTC 1711
Qy |||||
1441 QGTTCAAATTTATTAACCTCAATATTCATCTCAGAAGCATACAGGTGAAAAACCATACAAA 1500
Db |||||
1712 QGTTCAAATTTATTAACCTCAATATTCATCTCAGAAGCATACAGGTGAAAAACCATACAAA 1771
Qy |||||
1501 TGTGAAATTTTGTGAATATGCTGCAAGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA 1560
Db |||||
1772 TGTGAAATTTTGTGAATATGCTGCAAGCCAGAGACATCTCTGAGGTATCACTTTGGAGAGA 1831
Qy |||||
1561 CATCAAGGAAAACAAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG 1620
Db |||||
1832 CATCAAGGAAAACAAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGGTAAAAATCAG 1891
Qy |||||
1621 GACACTGAGATGCACATTAACCGCTGACAGTGCAGAAACCAAAAATTTTGAAGAATTT 1680
Db |||||
1892 GACACTGAGATGCACATTAACCGCTGACAGTGCAGAAACCAAAAATTTTGAAGAATTT 1951
Qy |||||
1681 TTTGATGCTGCAAGAGATTTACAGGAGTCCACCTGCAAAAGCAGCTTAAAGAGATGCT 1740
Db |||||
1952 TTTGATGCTGCAAGAGATTTACAGGAGTCCACCTGCAAAAGCAGCTTAAAGAGATGCT 2011
Qy |||||
1741 TCTGTTTTCAGAATGTTCTGGGAGCGCTGTCTCTCACAGCACACAAAGATATCTCAG 1800
Db |||||
2012 TCTGTTTTCAGAATGTTCTGGGAGCGCTGTCTCTCACAGCACACAAAGATATCTCAG 2071
Qy |||||
1801 GATTTCCATAAAATGCAAGTGCACAGTCTCATTAAGTGAATAAAAACCTTACCCCT 1860
Db |||||
2072 GATTTCCATAAAATGCAAGTGCACAGTCTCATTAAGTGAATAAAAACCTTACCCCT 2131
Qy |||||
1861 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATACCTCATC 1920
Db |||||
2132 GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATACCTCATC 2191
Qy |||||
1921 TGTAGAACCAAGGCGGATGTTACTCTCTCCGATGTCAGTACCAACCAATACCTTGA 1980
Db |||||
2192 TGTAGAACCAAGGCGGATGTTACTCTCTCCGATGTCAGTACCAACCAATACCTTGA 2251
Qy |||||
1981 GTTAGCCCCAAGAGAGAACCAAGCAGCTGACTGCAGATCAGGCAAGAGTGTG 2040
Db |||||
2252 GTTAGCCCCAAGAGAGAACCAAGCAGCTGACTGCAGATCAGGCAAGAGTGTG 2311
Qy |||||
2041 GATTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATTTGCCCGCAATT 2100
Db |||||
2312 GATTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATTTGCCCGCAATT 2371
Qy |||||
2101 TCTTTGAGTAAAGTTTGATTTCAAGATACCTGTCTCATTTGTACCTTCAAGACATTT 2160
Db |||||
2372 TCTTTGAGTAAAGTTTGATTTCAAGTATCACCTGTCTCATTTTGTACCTTCAAGACATTT 2431
Qy |||||
2161 TATCCAGAAGTTTAAATGATGCACAGAGACTGGAGCATAAATCAATCTCGACGTTCAAT 2220
Db |||||
2432 TATCCAGAAGTTTAAATGATGCACAGAGACTGGAGCATAAATCAATCTCGACGTTCAAT 2491
Qy |||||
2221 AAAAATCTCGAAACCAAGTCTTGTAGAAAGTGCAGTACCGGATGCCCGCAGCGTTG 2280
Db |||||
2492 AAAAATCTCGAAACCAAGTCTTGTAGAAAGTGCAGTACCGGATGCCCGCAGCGTTG 2551
Qy |||||
2281 CTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Db |||||
2552 CTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2611
Qy |||||
2341 CCGGCGCAGTCCAAATCTCTGCATCTCGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC 2400
Db |||||
2612 CCGGCGCAGTCCAAATCTCTGCATCTCGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC 2671
Qy |||||
2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2460
Db |||||
2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2731
Qy |||||
2461 CACAGACCAAGCAGAGAAATGTGGGGTCCAAAGGGCGGCCACAGGCAACAGCAATCTGAG 2520

Db 1112 AAGAAGCCTGTGATGTCATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGGCTTGGCAG 1171
Qy 901 CTGGCTACCAAGGAAAGATTGGCATTTGGCAAGAGTGAAGGAATCGGGCCNAGAAGG 960
Db 1172 CTGGCTACCAAGGAAAGATTGGCATTTGGCAAGAGTGAAGGAATCGGGCCNAGAAGG 1231
Qy 961 AGCACCGACAAACGACGATTTCGAGTTCGAGAGAGAGCTTGGAGAAACAAATAAGGGCAGT 1020
Db 1232 AGCACCGACAAACGACGATTTCGAGTTCGAGAGAGAGCTTGGAGAAACAAATAAGGGCAGT 1291
Qy 1021 TGTGAGGCTCTCTCGCAAGAGAAAGAGAGTGCAAAACACTCCACGGCGAAGCGCCCTCC 1080
Db 1292 TGTGAGGCTCTCTCGCAAGAGAAAGAGAGTGCAAAACACTCCACGGCGAAGCGCCCTCC 1351
Qy 1081 GTGACCGGGATCCNAGTTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGACCGGGATCCNAGTTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAAAGCTTTCAGAACTTACCACTACCAAGCTGGTCTTGGCACTCCAGGGTCCACAGAGGAC 1200
Db 1412 GGCAAAGCTTTCAGAACTTACCACTACCAAGCTGGTCTTGGCACTCCAGGGTCCACAGAGGAC 1471
Qy 1201 CGGAGGCGCGGCGGAGTGCGCCACCATGTCTGTGGAAGGAGAGCGCCGGGAGACGTGT 1260
Db 1472 CGGAGGCGCGGCGGAGTGCGCCACCATGTCTGTGGAAGGAGAGCGCCGGGAGACGTGT 1531
Qy 1261 TCTCTGACCTCGCGCCCTCTCGATGAAATGGAGCGGTGGATCGAGGGGAAAGGTGGT 1320
Db 1532 TCTCTGACCTCGCGCCCTCTCGATGAAATGGAGCGGTGGATCGAGGGGAAAGGTGGT 1591
Qy 1321 TCTGAAGACGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATATAAAATGATGAT 1380
Db 1592 TCTGAGACGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATATAAAATGATGAT 1651
Qy 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAGGTTTTTC 1440
Db 1652 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATGTGGAAGGTTTTTC 1711
Qy 1441 GGTTCAAATTTATACCTCAATATTTCATCTCAGAACGATACAGGTGAAAAACCATACAA 1500
Db 1712 GGTTCAAATTTATACCTCAATATTTCATCTCAGAACGATACAGGTGAAAAACCATACAA 1771
Qy 1501 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCTGAGGTATCACTTGGAGAGA 1831
Qy 1561 CATCAAGGAAAAACAAACCGATGTGCTGTAAGTCAAGAACGATGGTAAAAATCAG 1620
Db 1832 CATCAAGGAAAAACAAACCGATGTGCTGTAAGTCAAGAACGATGGTAAAAATCAG 1891
Qy 1621 GACACTGAAGATGCATTAACCGCTGACAGTGCCAACCAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAAGATGCATTAACCGCTGACAGTGCCAACCAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTGCCAAAGATTTACAGGCACTTCCAGTGCCTGCAAAAGCAGCTTAAGGAGATGCT 1740
Db 1952 TTTGATGTGCCAAAGATTTACAGGCACTTCCAGTGCCTGCAAAAGCAGCTTAAGGAGATGCT 2011
Qy 1741 TCTGTTTTTTCAGAAATGTTTGGGAGCGCTGCTCTCACAGCACAAAGATACTCAG 1800
Db 2012 TCTGTTTTTTCAGAAATGTTTGGGAGCGCTGCTCTCACAGCACAAAGATACTCAG 2071
Qy 1801 GATTTCCATAAAATGCAGCTGATGACAGTGTGATTAAGTGAATATAAAACCCCTACCCCT 1860
Db 2072 GATTTCCATAAAATGCAGCTGATGACAGTGTGATTAAGTGAATATAAAACCCCTACCCCT 2131
Qy 1861 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 1920
Db 2132 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTTGAAACTCAGGCAAAATAACCTCATC 2191
Qy 1921 TGTAGAACCAAGCGGATGTTTACTCTCTCTCCGATGGCAGTACCAACCATACCTTGAA 1980
Db 2192 TGTAGAACCAAGCGGATGTTTACTCTCTCTCCGATGGCAGTACCAACCATACCTTGAA 2251

Qy 1981 GTTAGCCCCAAGAGAGCAAAACGAGACCGCAGCTGACTGACAGATACAGGCCAAGTGTG 2040
Db 2252 GTTAGCCCCAAGAGAGCAAAACGAGACCGCAGCTGACTGACAGATACAGGCCAAGTGTG 2311
Qy 2041 GATTGTACAGAAACCTTTAAATTTATTCGTGGGGCTCTTCAAAATGTCGCCGCAATT 2100
Db 2312 GATTGTACAGAAACCTTTAAATTTATTCGTGGGGCTCTTCAAAATGTCGCCGCAATT 2371
Qy 2101 TCTTTGAGTAAAGTGTGATTTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAGTGTGATTTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAAATAAATCAATCTTCAAGTTCAT 2220
Db 2432 TATCCAGAAGTTTAAATGATGACACAGAGACTGGAGCATAAATAAATCAATCTTCAAGTTCAT 2491
Qy 2221 AAAAATCTGTCAAAACAAAGTCTTGTCTTAGAAGTCGACGTAACGGATGCCCGCAGCGTTG 2280
Db 2492 AAAAATCTGTCAAAACAAAGTCTTGTCTTAGAAGTCGACGTAACGGATGCCCGCAGCGTTG 2551
Qy 2281 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTC 2340
Db 2552 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTTC 2611
Qy 2341 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCGCAGGC 2400
Db 2612 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCGCAGGC 2671
Qy 2401 AAGGCCCTCTGACCTTCA - GGGATAGACTCTAGACATTTTAGCCCCAAAGTAACCTGAAGTC 2459
Db 2672 AAGGCCCTCTGACCTTCAAGGGGATAGACTCTAGACATTTTAGCCCCAAAGTAACCTGAAGTC 2731
Qy 2460 CCAAGACCAAGCAGAGATGTGGGGTCCAAAGGGCCGCCACAGGCAACAGCAATCTGA 2519
Db 2732 CCAAGACCAAGCAGAGATGTGGGGTCCAAAGGGCCGCCACAGGCAACAGCAATCTGA 2791
Qy 2520 GATGTTTCTTAAACACAGTGTTCCTCTGCAACCGGATAAGACAAAGAACCCGAGACAAA 2579
Db 2792 GATGTTTCTTAAACACAGTGTTCCTCTGCAACCGGATAAGACAAAGAACCCGAGACAAA 2851
Qy 2580 ATTGAACCTCTTCCAGTAGTCTCTTTCAGGCCACCTCGGAGCAGTAACATCAATGG 2639
Db 2852 ATTGAACCTCTTCCAGTAGTCTCTTTCAGGCCACCTCGGAGCAGTAACATCAATGG 2911
Qy 2640 TTCCATCAGCTACCCCGCCAAAGACAGACGCCCTGGGCACCTCCGGGAAGAGACTATT 2699
Db 2912 TTCCATCAGCTACCCCGCCAAAGACAGACGCCCTGGGCACCTCCGGGAAGAGACTATT 2971
Qy 2700 CTGTAATCGGAGTGCAGCAATACTGCAAGCAAGATTTTGGTAGCCCCCTTCCAAAAGACT 2759
Db 2972 CTGTAATCGGAGTGCAGCAATACTGCAAGCAAGATTTTGGTAGCCCCCTTCCAAAAGACT 3031
Qy 2760 GAAGTCCAGCGTGTGTCCTTTGAGCGTTGACAGCCCGGGCCAAATTAAGAGAGGCTA 2819
Db 3032 GAAGTCCAGCGTGTGTCCTTTGAGCGTTGACAGCCCGGGCCAAATTAAGAGAGGCTA 3091
Qy 2820 TGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCTGTTTACCGCAGACTGTGT 2879
Db 3092 TGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCTGTTTACCGCAGACTGTGT 3151
Qy 2880 GTATCCGTCCGAGGCGCTCCCTCCCAACCAAGGTTTCTGAGCTCCAGCAGGTGATTC 2939
Db 3152 GTATCCGTCCGAGGCGCTCCCTCCCAACCAAGGTTTCTGAGCTCCAGCAGGTGATTC 3211
Qy 2940 TCCAAATGTGTGACTGTTCAGAACCCCTATGTTGGCTCCGGGCCACTTTTACACTGTGT 2999
Db 3212 TCCAAATGTGTGACTGTTCAGAACCCCTATGTTGGCTCCGGGCCACTTTTACACTGTGT 3271
Qy 3000 GCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAG - ----- 3034
Db 3272 GCCTGCTGGTAGTCCAGCATCCAGCTCGACGTTAGAGGAAAGGCGCTGTGTATATCA 3331

QY	3035	-----	3034
Db	3332	ACACTTATCTAACAGCATGGCACAAGAGAACTATGAGAAATTTATTTGGGAATGCACA	3391
QY	3035	-----	3046
Db	3392	TTATCGACCAATGACACAAAAAACTTGATTCACCTAAATAGGGGAAAAAGGCTCTGGTG	3451
QY	3047	GATGTCTAGTGCTTACTCTCCCATGAAATTTAAATTTTACTTTCATCTTTTGAGAACGGAATGG	3106
Db	3452	GATGTCTAGTGCTTACTCTCCCATGAAATTTAAATTTTACTTTCATCTTTTGAGAACGGAATGG	3511
QY	3107	TGAAGCTACTGAAATGAAGCTGCTGATTTGTTACTGTACATAAACAATAGAGGAATCTGCAG	3166
Db	3512	TGAAGCTACTGAAATGAAGCTGCTGATTTGTTACTGTACATAAACAATAGAGGAATCTGCAG	3571
QY	3167	GGAACTACTACAGTTGTGTAA	3186
Db	3572	GGAACTACTACAGTTGTGTAA	3591
RESULT 5			
AL157838/c			
LOCUS	AL157838	128871 bp	DNA linear PRI 23-FEB-2001
DEFINITION	Human DNA sequence from clone RP4-724E16 on chromosome 20q13.12-13.32 Contains the ZNF217 gene for zinc finger protein 217, a novel gene, a putative novel gene, ESTs, GSSs, STSs and two CpG islands, complete sequence.		
ACCESSION	AL157838		
VERSION	AL157838.24	GI:9588158	
KEYWORDS	HTG; CpG island; zinc finger; ZNF217.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 128871)		
TITLE	Direct Submission		
JOURNAL	Submitted (23-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk		
COMMENT	requests: clonerequest@sanger.ac.uk On Jul 31, 2000 this sequence version replaced gi:9408255. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: En: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr20 This sequence is the entire insert of clone RP4-724E16 The true right end of clone RP5-823G15 is at 19684 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-724E16 is from the library RPC1-4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2.		
FEATURES	Location/Qualifiers		
source	1. 128871 /organism="Homo sapiens"		

```
repeat_region /note="MERSA repeat: matches 4. .183 of consensus"
9980. .10108
repeat_region /note="L12 repeat: matches 2622. .2747 of consensus"
10128. 10437
repeat_region /note="AluX repeat: matches 2. .312 of consensus"
10654. .10724
repeat_region /note="MIR repeat: matches 63. .137 of consensus"
10781. .11083
repeat_region /note="AluX repeat: matches 1. .306 of consensus"
complement(10897. .11319)
/note="match: GSS: Em:AQ499673"
complement(10991. .11319)
/note="match: GSS: Em:AQ283579"
11462. .11759
repeat_region /note="AluX repeat: matches 1. .299 of consensus"
11878. .12182
repeat_region /note="AluX repeat: matches 1. .300 of consensus"
12183. .12587
repeat_region /note="L1PA8 repeat: matches 5757. .6163 of consensus"
13333. .13637
repeat_region /note="MLTIG repeat: matches 3. .338 of consensus"
14201. .14507
repeat_region /note="AluY repeat: matches 1. .308 of consensus"
14540. .14694
repeat_region /note="L12 repeat: matches 2550. .2709 of consensus"
14764. .14868
repeat_region /note="MIR repeat: matches 103. .208 of consensus"
14891. .15152
repeat_region /note="AluJo repeat: matches 12. .299 of consensus"
15162. .15293
repeat_region /note="L12 repeat: matches 1697. .1824 of consensus"
15317. .15544
repeat_region /note="L1MB3 repeat: matches 5943. .6173 of consensus"
16306. .16603
repeat_region /note="AluSc repeat: matches 1. .306 of consensus"
16610. .16901
repeat_region /note="AluX repeat: matches 1. .292 of consensus"
16989. .17035
repeat_region /note="MLTIC repeat: matches 420. .463 of consensus"
17036. .17338
repeat_region /note="AluJo repeat: matches 1. .301 of consensus"
17339. .17730
repeat_region /note="MLTIC repeat: matches 1. .420 of consensus"
17958. .18269
repeat_region /note="AluX repeat: matches 1. .311 of consensus"
18288. .18315
repeat_region /note="MERS1 repeat: matches 32. .59 of consensus"
18383. .18507
repeat_region /note="MIR repeat: matches 23. .148 of consensus"
18544. .18650
repeat_region /note="MIR repeat: matches 137. .261 of consensus"
18799. .18820
repeat_region /note="11 copies 2 mer aa 100% conserved"
18835. .19184
repeat_region /note="THE1A repeat: matches 1. .354 of consensus"
19231. .19522
repeat_region /note="match: STS: Em:HS153YF5"
19297. .19346
repeat_region /note="25 copies 2 mer ca 100% conserved"
19426. .19564
repeat_region /note="MIR repeat: matches 6. .147 of consensus"
19624. .19916
repeat_region /note="AluX repeat: matches 1. .295 of consensus"
20005. .20204
repeat_region /note="MER20 repeat: matches 1. .218 of consensus"
20215. .20317
repeat_region /note="MERSA repeat: matches 4. .108 of consensus"
20410. .20541
repeat_region /note="MERSB repeat: matches 1. .120 of consensus"
20544. .20687
repeat_region /note="AluSx/x repeat: matches 154. .297 of consensus"
20689. .20814
repeat_region /note="FLAM_A repeat: matches 14. .136 of consensus"
```

```
repeat_region 20816. .21218
/note="Charlie4a repeat: matches 17. .427 of consensus"
repeat_region 21337. .21634
/note="AluY repeat: matches 1. .297 of consensus"

Query Match 49.0%; Score 1561.2; DB 9; Length 128871;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1482 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATAGTCGACGCCAGAGACATCTCT 1541
Db |||||
Qy 65718 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATAGTCGACGCCAGAGACATCTCT 65659
Db |||||
Qy 1542 GAGGTATCACTTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 1601
Db |||||
Qy 65858 GAGGTATCACTTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGCTGAAGTCAA 65599
Db |||||
Qy 1602 GAACGATGTTAAATCAGGACATGAAGATGCTATTAACCCGCTGACAGCTGGCGAAAC 1661
Db |||||
Qy 65598 GAACGATGTTAAATCAGGACATGAAGATGCTATTAACCCGCTGACAGCTGGCGAAAC 65539
Db |||||
Qy 1662 CAAAAATTTGAAAAGATTTTGTGATGTCGCAAGATGTTACAGSCAGTCCACCTGCAAA 1721
Db |||||
Qy 65538 CAAAAATTTGAAAAGATTTTGTGATGTCGCAAGATGTTACAGSCAGTCCACCTGCAAA 65479
Db |||||
Qy 1722 GCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTTGGGCGAGCGCTGCTCTCAC 1781
Db |||||
Qy 65478 GCAGCTTAAGGAGATGCTCTGTTTTTCAGAAATGTTCTTGGGCGAGCGCTGCTCTCAC 65419
Db |||||
Qy 1782 AGCACAAGATGACTCAGGATTTTCATATAAATGCGAGCTGATGACAGTGTGATAAGT 1841
Db |||||
Qy 65418 AGCACAAGATGACTCAGGATTTTCATATAAATGCGAGCTGATGACAGTGTGATAAGT 65359
Db |||||
Qy 1842 GAATATAAACCTACCTGCTTACCTGAGCTGTTAAAAAGAGATCAGCAGTTGAAAC 1901
Db |||||
Qy 65358 GAATATAAACCTACCTGCTTACCTGAGCTGTTAAAAAGAGATCAGCAGTTGAAAC 65299
Db |||||
Qy 1902 TCAGGCAAAATAACCTCATCTGTAGAAACCAAGCGCGATGTTACTCCTCTCCGGATGGCAG 1961
Db |||||
Qy 65298 TCAGGCAAAATAACCTCATCTGTAGAAACCAAGCGCGATGTTACTCCTCTCCGGATGGCAG 65239
Db |||||
Qy 1962 TACCACCCATTAACCTTGAAGTTAGCCCAAGAGAGAAACCGAGACCGCAGCTGACTG 2021
Db |||||
Qy 65238 TACCACCCATTAACCTTGAAGTTAGCCCAAGAGAGAAACCGAGACCGCAGCTGACTG 65179
Db |||||
Qy 2022 CAGATACAGGCCAAGTGTGAGTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
Db |||||
Qy 65178 CAGATACAGGCCAAGTGTGAGTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 65119
Db |||||
Qy 2082 TCACAAATTTGCCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACCTGTCCATT 2141
Db |||||
Qy 65118 TCACAAATTTGCCCGGCAATTTCTTTGAGTAAAGTTTGATTTCCAAAGTATCACCTGTCCATT 65059
Db |||||
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTATGATGACCCAGAGACCTGGAGCATAA 2201
Db |||||
Qy 65058 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTATGATGACCCAGAGACCTGGAGCATAA 64999
Db |||||
Qy 2202 ATACAATCCTGACGTTTCATATAAAGTGTGCAAAACAAAGTCTTTAGAGTTCGACGTAC 2261
Db |||||
Qy 64998 ATACAATCCTGACGTTTCATATAAAGTGTGCAAAACAAAGTCTTTAGAGTTCGACGTAC 64939
Db |||||
Qy 2262 CGGATGCCCGCGCAGCTTGTGGGAAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db |||||
Qy 64938 CGGATGCCCGCGCAGCTTGTGGGAAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 64879
Db |||||
Qy 2322 CAAGCCCAAGTCTGTTTCCCGGCGCAGTCAAATCCCTGCCATCTGGAAGGGAAGCA 2381
Db |||||
Qy 64878 CAAGCCCAAGTCTGTTTCCCGGCGCAGTCAAATCCCTGCCATCTGGAAGGGAAGCA 64819
Db |||||
Qy 2382 GAGCCCTCCTGGCCAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGACCTTTAGC 2441
Db |||||
Qy 64818 GAGCCCTCCTGGCCAGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGACCTTTAGC 64759
Db |||||
```



```
QY 2442 CCCAAGTAACCTGAAGTCCACAGACCACAGCAGAAATGGGGGTCCAAAGGGCGCCAC 2501
Db 64758 CCCAAGTAACCTGAAGTCCACAGACCACAGCAGAAATGGGGGTCCAAAGGGCGCCAC 64699
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAGTGTTCCTCCCTGCACACCGGATAAGAC 2561
Db 64698 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAGTGTTCCTCCCTGCACACCGGATAAGAC 64639
QY 2562 AAAAAGACCCGAGACAAAATTGAAACCTCTTCAGTAGTCTCTTCAGCCACACCTCGG 2621
Db 64638 AAAAAGACCCGAGACAAAATTGAAACCTCTTCAGTAGTCTCTTCAGCCACACCTCGG 64579
QY 2622 CAGCAGTAACATCAATGTTTCCATCGACTACCCCGCAAGAACGACAGCCGTTGGGCACC 2681
Db 64578 CAGCAGTAACATCAATGTTTCCATCGACTACCCCGCAAGAACGACAGCCGTTGGGCACC 64519
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATCTGCAGCAGAAATTTGGTGA 2741
Db 64518 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATCTGCAGCAGAAATTTGGTGA 64459
QY 2742 GCCCCTTCCAAAAGACTGAAGTCAGCGTGTTCGCTTACGTTGACAGCCCGGGGC 2801
Db 64458 GCCCCTTCCAAAAGACTGAAGTCAGCGTGTTCGCTTACGTTGACAGCCCGGGGC 64399
QY 2802 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 2861
Db 64398 CAATTACAGAAGAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT 64339
QY 2862 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCGTGCCTCCCAACCAAGGTTCTCTGAG 2921
Db 64338 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCGTGCCTCCCAACCAAGGTTCTCTGAG 64279
QY 2922 TTCAGCAGAGTGCATTTCCAAATGTGCTGACTGTTCCAGAGCCCTATGFGGCTCCGG 2981
Db 64278 TTCAGCAGAGTGCATTTCCAAATGTGCTGACTGTTCCAGAGCCCTATGFGGCTCCGG 64219
QY 2982 GCCACTTTACACTTGTGCTGCTGCTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 3041
Db 64218 GCCACTTTACACTTGTGCTGCTGCTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 64159
QY 3042 TGGTGGATGTCAGT 3055
Db 64158 TGCATGAGGGCGT 64145

RESULT 6
BD085733
LOCUS BD085733 10365 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085733
VERSION BD085733.1 GI:22631343
KEYWORDS JP 2001524802-A/9.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 10365)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 9 04-DEC-2001;
COMMENT THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/9
PD 04-DEC-2001
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/580395,16-OCT-1996 US 08/731499 PR
PI 17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11, C12O1/68, A61K48/00
CC Description of Artificial Sequence: Genomic Sequence encoding
ZABCI
```

```
FEATURES
source
FH Key Location/Qualifiers
FT source 1..10365
/organism='Artificial Sequence'
Location/Qualifiers
1..10365
/organism='synthetic construct'
/mol_type='genomic DNA'
/db_xref='taxon:32630'
ORIGIN
Query Match 49.0%; Score 1559.6; DB 6; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1482 AGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 1541
Db 8286 AGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGACATCTCT 8345
QY 1542 GAGGTATCACCTTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 1601
Db 8346 GAGGTATCACCTTGGAGAGACATCAAGGAAACCAACCCGATGTTGCTGTAAGTCAA 8405
QY 1602 GAAAGATCGTAAAAATCAGGACACTGAAGATGCACATTAACCCGCTGACAGTGCAGCAAC 1661
Db 8406 GAAAGATCGTAAAAATCAGGACACTGAAGATGCACATTAACCCGCTGACAGTGCAGCAAC 8465
QY 1662 CAAAAATTTGAAAGATTTTGTGATGGTGCACCAAGATGTTACAGCAGTCCACCTGCACA 1721
Db 8466 CAAAAATTTGAAAGATTTTGTGATGGTGCACCAAGATGTTACAGCAGTCCACCTGCACA 8525
QY 1722 GCAGCTTAAGAGATGCCCTCTGTTTTTTCAGAAATGTTCTGGGCAGCGCTGCTCTCACC 1781
Db 8526 GCAGCTTAAGAGATGCCCTCTGTTTTTTCAGAAATGTTCTGGGCAGCGCTGCTCTCACC 8585
QY 1782 AGCACACAAAGATACTCAGGATTTCCATAAAATGCAGTGAATGACAGTGTGATAAAGT 1841
Db 8586 AGCACACAAAGATACTCAGGATTTCCATAAAATGCAGTGAATGACAGTGTGATAAAGT 8645
QY 1842 GAATAAAACCCCTACCCCTGCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8646 GAATAAAACCCCTACCCCTGCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTTGAAAC 8705
QY 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 1961
Db 8706 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 8765
QY 1962 TACCACCATTAACCTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 2021
Db 8766 TACCACCATTAACCTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 8825
QY 2022 CAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 2081
Db 8826 CAGATACAGGCCAAGTGTGATTTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 8885
QY 2082 TCACAAATGCCCGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATT 2141
Db 8886 TCACAAATGCCCGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATT 8945
QY 2142 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGACATGGAGCATAA 2201
Db 8946 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGACATGGAGCATAA 9005
QY 2202 ATACAATCCTGACGTTTCAAAAACTGTGCGAAAAAAGTCTTGTCTTAGAAGTTCGACGTAC 2261
Db 9006 ATACAATCCTGACGTTTCAAAAACTGTGCGAAAAAAGTCTTGTCTTAGAAGTTCGACGTAC 9065
QY 2262 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCCCTCCCTCTCTAGTTCTGTAAACC 2321
Db 9066 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCCCTCCCTCTCTAGTTCTGTAAACC 9125
QY 2322 CAAGCCCAAGTCTGCTTTCCCGCGCGCAGTCCAAATTCCTGCGCATCTCGAGAGGGAAGCA 2381
Db 9126 CAAGCCCAAGTCTGCTTTCCCGCGCGCAGTCCAAATTCCTGCGCATCTCGAGAGGGAAGCA 9185
```


Qy	2382	GAGCCCTCTCTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC	2441	JOURNAL	Submitted (10-OCT-2000) Cancer Genetics, UCSF Cancer Center, 2340
Db	9186	GAGCCCTCTCTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC	9245	FEATURES	Sutter St., Rm. S151, San Francisco, CA 94706, USA
Qy	2442	CCCAAGTAACCTGAAGTCCCAAGACACAGACAGATGTGGGGTCCAGGGGGCGCCAC	2501	source	1. 121143
Db	9246	CCCAAGTAACCTGAAGTCCCAAGACACAGACAGATGTGGGGTCCAGGGGGCGCCAC	9305		/organism="Homo sapiens"
Qy	2502	CAGCAACAGCAATCTGAGATGTTTCTAAACACAGTCTTCCCTGACCGGATAGAC	2561		/mol_type="genomic DNA"
Db	9306	CAGCAACAGCAATCTGAGATGTTTCTAAACACAGTCTTCCCTGACCGGATAGAC	9365		/db_xref="taxon:9606"
Qy	2562	AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCTCTCAGCCACCCCTCGG	2621		/chromosome="20"
Db	9366	AAAAAGACCCGAGACAAAATTGAAACCTCTTCCAGTAGCTCTCTCAGCCACCCCTCGG	9425		/map="20q13.2"
Qy	2622	CAGCAGTAACATCAATGTTTCATCGACTACCCCGCCGAAGACAGACGCCGTGGGCACC	2681	misc_feature	/clones="97: 127: LBNL h117 and h79"
Db	9426	CAGCAGTAACATCAATGTTTCATCGACTACCCCGCCGAAGACAGACGCCGTGGGCACC	9485		/note="isolated in Collins lab and sequenced by the Genome Center at LBNL"
Qy	2682	TCCGGAGAGACATATTTCTGTAATCGGAGTCCAGCAATCTCAGCAGAAATTTGGTGA	2741	repeat_region	complement(1..66093)
Db	9486	TCCGGAGAGACATATTTCTGTAATCGGAGTCCAGCAATCTCAGCAGAAATTTGGTGA	9545		/note="overlaps Homo sapiens chromosome 20 clone RPS-823G15 deposited in GenBank Accession Number AL354993.19"
Qy	2742	GCCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACCTTGACCGCCCGGGGC	2801	repeat_region	194..468
Db	9546	GCCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACCGCCCGGGGC	9605		/rpt_family="Alu"
Qy	2802	CAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGCATCATCACT	2861	repeat_region	/rpt_type=dispersed
Db	9606	CAATACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGCATCATCACT	9665		complement(513..653)
Qy	2862	GTTACCGCAGACTGTGTATCCGTCCAGCGCTGCCTCCCAACCAAGGTTCCCTGAG	2921	repeat_region	/rpt_family="MER5"
Db	9666	GTTACCGCAGACTGTGTATCCGTCCAGCGCTGCCTCCCAACCAAGGTTCCCTGAG	9725		/rpt_type=dispersed
Qy	2922	CTCCAGCCAGCTCGAATCTCCAAATGTCTGACTGTTCCAGAGCCCTATGTTGGTCCGG	2981	repeat_region	1053..1153
Db	9726	CTCCAGCCAGCTCGAATCTCCAAATGTCTGACTGTTCCAGAGCCCTATGTTGGTCCGG	9785		/rpt_family="MER20"
Qy	2982	GCCACTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3041	repeat_region	/rpt_type=dispersed
Db	9786	GCCACTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	9845		complement(1356..1676)
Qy	3042	TGGTGGATGTCAGT 3055		repeat_region	/rpt_family="Alu"
Db	9846	TGCATGAGGGCGT 9859			/rpt_type=dispersed
RESULT 7				repeat_region	2594..2879
AF312915/c					/rpt_family="Alu"
LOCUS	AF312915	121143 bp	DNA	linear	PR1 13-JUN-2001
DEFINITION	Homo sapiens chromosome 20 clones 97 and 127, complete sequence.				complement(2896..4437)
ACCESSION	AF312915				/note="similar to flow-sorted chromosome 20 HindIII fragment, SC20pF17C6 deposited in GenBank Accession Number Z94682.1"
VERSION	AF312915.1	GI:11094030			3106..3412
KEYWORDS					/rpt_family="Alu"
SOURCE	Homo sapiens (human)				/rpt_type=dispersed
ORGANISM	Homo sapiens				3889..4206
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				/rpt_family="Alu"
AUTHORS	1 (bases 1 to 121143) Collins, C., Volik, S., Kowbel, D., Ginzinger, D., Ylstra, B., Cloutier, T., Hawkins, T., Predki, P., Martin, C., Wernick, M., Kuo, W.L., Alberts, A. and Gray, J.W.				4449..4764
TITLE	Comprehensive genome sequence analysis of a breast cancer amplicon				/rpt_type=dispersed
JOURNAL	Genome Res. 11 (6), 1034-1042 (2001)				complement(5671..5987)
MEDLINE	21275464				/rpt_family="Alu"
REFERENCE	11381030				/rpt_type=dispersed
AUTHORS	Volik, S., Collins, C., Gray, J., Wernick, M., Kowbel, D., Stultz, K. and Martin, C.				6723..7004
TITLE	Direct Submission				/rpt_family="Alu"

```
/product="hypothetical"
/note="similar to Homo sapiens hypothetical protein
OVC10-2 deposited in GenBank Accession Numbers AF230201.1,
AF000365 and in NM_018692.1"
/evidence=not experimental
repeat_region 15301..15588
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature 16873..17073
/note="similar to Homo sapiens STS WI-16697 deposited in
GenBank Accession Number G21490.1"
repeat_region 17577..17857
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(19034..19333)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20136..20719
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20776..21071
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22344..22638)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 22726..22784
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(22984..23268)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(23349..23646)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(24008..24189)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region complement(24502..24795)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 24883..24941
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(25140..25425)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(25507..25805)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(26170..26352)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region 26841..27129
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 27270..33290
/rpt_family="I1"
/rpt_type=dispersed
repeat_region 28512..28730
/rpt_family="MER25"
/rpt_type=dispersed
repeat_region 34498..35009
/rpt_family="MLT2B2"
/rpt_type=dispersed
repeat_region 35035..35332
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 35343..35522
/rpt_family="Alu"
/rpt_type=dispersed
misc_feature 35593..121143
/note="overlaps Homo sapiens clone RP4-724E16 on
```

```
chromosome 20 deposited in GenBank Accession Number
AL157838.24"
repeat_region 36137..36250
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 36366..36690
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 36949..37020
/rpt_family="MERS"
/rpt_type=dispersed
repeat_region complement(37692..37766)
/rpt_family="MERS"
/rpt_type=dispersed
repeat_region complement(38006..38276)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 38863..38984
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(39791..40072)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(40100..40395)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 40642..41118
/rpt_family="MER1"
/rpt_type=dispersed
repeat_region complement(41228..41348)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 41649..41938
Query Match 49.0%; Score 1559.6; DB 9; Length 121143;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1482 AGGTGAAAAACCATCAAAATGTGAATTTTGTGAATATCTCTGAGCCCAAGACATCTCT 1541
Db 99470 AGGTGMAAAACCATCAAAATGTGAATTTTGTGAATATCTCTGAGCCCAAGACATCTCT 99411
QY 1542 GAGGTATCATTGGAGAGACATCACAGGAAAAAACAACCCGATGTTGCTGAAGTCNA 1601
Db 99410 GAGGTATCATTGGAGAGACATCACAGGAAAAAACAACCCGATGTTGCTGAAGTCNA 99351
QY 1602 GAACGATGCTAAAAATCAGGACACTGAAGATGTCACATATTAAACCGCTGACAGTCCCAAC 1661
Db 99350 GAACGATGCTAAAAATCAGGACACTGAAGATGTCACATATTAAACCGCTGACAGTCCCAAC 99291
QY 1662 CAAAAATTTGAAAAGATTTTTTGTGATGTCGCAAGATGTTACAGCAGTCCACCTGCAAA 1721
Db 99290 CAAAAATTTGAAAAGATTTTTTGTGATGTCGCAAGATGTTACAGCAGTCCACCTGCAAA 99231
QY 1722 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAAATGTTCTGGCAGCGCTGCTCTCACC 1781
Db 99230 GCAGCTTAAGGAGATGCCTTCTGTTTTTCAGAAATGTTCTGGCAGCGCTGCTCTCACC 99171
QY 1782 AGCACAAAAAGATCTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
Db 99170 AGCACAAAAAGATCTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 99111
QY 1842 GAATAAAAAACCTTACCCCTGCTTACCTGGACCTGTGTAATAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 99110 GAATAAAAAACCTTACCCCTGCTTACCTGGACCTGTGTAATAAAAAAGAGATCAGCAGTTGAAAC 99051
QY 1902 TCAGGCAATAAACCTCATCTCTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAG 1961
Db 99050 TCAGGCAATAAACCTCATCTCTAGAACCAAGCGCGATGTTACTCTCTCTCGGATGGCAG 98991
QY 1962 TACCACCCATACCTTTGAAGTTAGCCCCCAAGAGAGAACCAAGCGGAGCCGAGCTGACTG 2021
Db 98990 TACCACCCATACCTTTGAAGTTAGCCCCCAAGAGAGAACCAAGCGGAGCCGAGCTGACTG 98931
```

```
Qy 2022 CAGATACAGGCGCAAGTGTGGATTGTTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
Db |||
Qy 98930 CAGATACAGGCGCAAGTGTGGATTGTTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 98871
Db |||
Qy 2082 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTCACAAATGATCACCTGTCCATT 2141
Db |||
Qy 98870 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGATTCACAAATGATCACCTGTCCATT 98811
Db |||
Qy 2142 TTGTACCTTCAGACATTTTATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAA 2201
Db |||
Qy 98810 TTGTACCTTCAGACATTTTATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAA 98751
Db |||
Qy 2202 ATACAATCCTGACGTTTCATAAAAACTGCGAAAAAGTCTTGTCTTGAAGTGCACGTAC 2261
Db |||
Qy 98750 ATACAATCCTGACGTTTCATAAAAACTGCGAAAAAGTCTTGTCTTGAAGTGCACGTAC 98691
Db |||
Qy 2262 CGGATGCCCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTCTGTAAACC 2321
Db |||
Qy 98690 CGGATGCCCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTCTGTAAACC 98631
Db |||
Qy 2322 CAAGCCCAAGTCTGCTTTCCGCGCGAGTCCAAATCCCTGCATCTGCGAAGGGAGCA 2381
Db |||
Qy 98630 CAAGCCCAAGTCTGCTTTCCGCGCGAGTCCAAATCCCTGCATCTGCGAAGGGAGCA 98571
Db |||
Qy 2382 GAGCCCTCCTGGCCAGGCAAGGCCCTCTGCACTTCAGGGATAGACTCTAGCATTAGC 2441
Db |||
Qy 98570 GAGCCCTCCTGGCCAGGCAAGGCCCTCTGCACTTCAGGGATAGACTCTAGCATTAGC 98511
Db |||
Qy 2442 CCCAAGTAACTGAAGTCCCAAGACACAGAGATGTGGGGTCCAAAGGGCGGCCAC 2501
Db |||
Qy 98510 CCCAAGTAACTGAAGTCCCAAGACACAGAGATGTGGGGTCCAAAGGGCGGCCAC 98451
Db |||
Qy 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGGATAGAC 2561
Db |||
Qy 98450 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTCCCTGACCGGATAGAC 98391
Db |||
Qy 2562 AAAAAGACCCAGACAAAATTTGAAACCTCTTCCAGTAGCTCTTCTCAGCCCAACCTCGG 2621
Db |||
Qy 98390 AAAAAGACCCAGACAAAATTTGAAACCTCTTCCAGTAGCTCTTCTCAGCCCAACCTCGG 98331
Db |||
Qy 2622 CAGCAGTAACTGAATGTTCCATCGACTACCCCGCCAAAGAACGACGCCGTGGGCACC 2681
Db |||
Qy 98330 CAGCAGTAACTGAATGTTCCATCGACTACCCCGCCAAAGAACGACGCCGTGGGCACC 98271
Db |||
Qy 2682 TCCGGGAAGAGACTATTTCTGTAATCCGAGTGCAGCAATCTGCGAGCAATTTGGTGA 2741
Db |||
Qy 98270 TCCGGGAAGAGACTATTTCTGTAATCCGAGTGCAGCAATCTGCGAGCAATTTGGTGA 98211
Db |||
Qy 2742 GCGCCCTTCCAAAAAGACTGAAGTGCAGCGTGGTTCGCCCTTGACGTTGACCGCCGGGGC 2801
Db |||
Qy 98210 GCGCCCTTCCAAAAAGACTGAAGTGCAGCGTGGTTCGCCCTTGACGTTGACCGCCGGGGC 98151
Db |||
Qy 2802 CAATPACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGCATCAATCACT 2861
Db |||
Qy 98150 CAATPACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGCATCAATCACT 98091
Db |||
Qy 2862 GTTACCGCAGGACTGTGTATTCGTCGCGCGGCTGCCCTCCCAACCAAGTTCCTGAG 2921
Db |||
Qy 98090 GTTACCGCAGGACTGTGTATTCGTCGCGCGGCTGCCCTCCCAACCAAGTTCCTGAG 98031
Db |||
Qy 2922 CTCACGCGAGGTTCGATTTCTCAAAATGCTGACTGTTCCAGAGCCCTATGTTGGTCCGG 2981
Db |||
Qy 98030 CTCACGCGAGGTTCGATTTCTCAAAATGCTGACTGTTCCAGAGCCCTATGTTGGTCCGG 97971
Db |||
Qy 2982 GCCACTTTACCTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGCTTAGAGGTCT 3041
Db |||
Qy 97970 GCCACTTTACCTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGCTTAGAGGTAT 97911
Db |||
Qy 3042 TGGTGGATGTCAGT 3055
Db |||
Qy 97910 TGCATGAGGGGCGT 97897
Db |||
```

```
RESULT 8
AC116668 105023 bp DNA linear HTG 01-MAY-2002
LOCUS Trypanosoma brucei chromosome V clone RPC193-45E22, *** SEQUENCING
DEFINITION IN PROGRESS ***, 3 unordered pieces.
AC116668
AC116668.5 GI:20376999
VERSION HTG; HTGS PHASE1.
KEYWORDS Trypanosoma brucei
SOURCE Trypanosoma brucei
ORGANISM Trypanosoma brucei
Eukaryota; Euzoenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE
1 (bases 1 to 105023)
El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
Petersen,J., Hou,L., Zhao,H., Mason,T., Militscher,J., Pai,G., Van
Alen,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.,
Trypanosoma brucei GUTat10.1 RPC193-45E22 BAC genomic sequence
Unpublished
2 (bases 1 to 105023)
El-Sayed,N.M., Khalak,H. and Adams,M.D.
Direct Submission
Submitted (02-APR-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
3 (bases 1 to 105023)
El-Sayed,N.M., Khalak,H. and Adams,M.D.
Direct Submission
Submitted (01-MAY-2002) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA
On May 1, 2002 this sequence version replaced gi:20340472.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 85781: contig of 85781 bp in length
* 85782 85806: gap of unknown length
* 85807 102956: contig of 17150 bp in length
* 102957 102981: gap of unknown length
* 102982 105023: contig of 2042 bp in length.
FEATURES
Location/Qualifiers
source
1..105023
/organism="Trypanosoma brucei"
/mol_type="genomic DNA"
/isolate="GUTat10.1"
/db_xref="taxon:5691"
/chromosome="V"
/clone="RPC193-45E22"
ORIGIN
Query Match 48.9%; Score 1556.4; DB 2; Length 105023;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1563; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT 1541
Db 12569 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT 12628
Qy 1542 GAGGTATCATCTGGAGAGACATCAAGAGAAAAACCAACCGATGTTGCTGAAGTCAA 1601
Db 12629 GAGGTATCATCTGGAGAGACATCAAGAGAAAAACCAACCGATGTTGCTGAAGTCAA 12688
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGCAAC 1661
Db 12689 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGCAAC 12748
Qy 1662 CAAAAATTTGAAAGATTTTGTGATGTCGCAAGATGTTTACAGCAGTCCACCTGCAAA 1721
Db 12749 CAAAAATTTGAAAGATTTTGTGATGTCGCAAGATGTTTACAGCAGTCCACCTGCAAA 12808
```

```
QY 1722 GCAGCTTAAGGAGATGCGCTTCTGTTTTCAGAAATGTTCTGGCAGGCGTGTCTCTCAC 1781
Db 12809 GCAGCTTAAGGAGATGCGCTTCTGTTTTCAGAAATGTTCTGGCAGGCGTGTCTCTCAC 12868
QY 1782 AGCACACAAAGATCTCAGGATTTCCATAAAATGAGCTGATGACAGTGTCTGATAAAGT 1841
Db 12869 AGCACACAAAGATCTCAGGATTTCCATAAAATGAGCTGATGACAGTGTCTGATAAAGT 12928
QY 1842 GNAATAAAACCTTACCCCTGCTTACTGGACCTGTGTTAAAAAGAGATCAGCAATTTGAAAC 1901
Db 12929 GAAATAAAACCTTACCCCTGCTTACTGGACCTGTGTTAAAAAGAGATCAGCAATTTGAAAC 12988
QY 1902 TCAGCAAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
Db 12989 TCAGCAAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 13048
QY 1962 TACCACCCATACCTTTGAAGTTAGCCCAAGAGAAAGCAAGCGAGACCGAGCTGACTG 2021
Db 13049 TACCACCCATACCTTTGAAGTTAGCCCAAGAGAAAGCAAGCGAGACCGAGCTGACTG 13108
QY 2022 CAGATACAGGCAAGTGTGGATTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 2081
Db 13109 CAGATACAGGCAAGTGTGGATTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCT 13168
QY 2082 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGTATCCCAAGTATCACTGTCCATT 2141
Db 13169 TCACAAATGCGCGCAATTTCTTTGAGTAAAGTTTGTATCCCAAGTATCACTGTCCATT 13228
QY 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAA 2201
Db 13229 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAGAGACTGGAGCATAA 13288
QY 2202 ATACAAATCCTGAGCTTCAATAAAACCTGCGAAACCAAGTCTTGTGTTAGAAATCGACGAC 2261
Db 13289 ATACAAATCCTGAGCTTCAATAAAACCTGCGAAACCAAGTCTTGTGTTAGAAATCGACGAC 13348
QY 2262 CGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db 13349 CGGATGCCCGCAGCGTGTCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 13408
QY 2322 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGATCTCGAAAGGGGAAGCA 2381
Db 13409 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCGATCTCGAAAGGGGAAGCA 13468
QY 2382 GAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGC 2441
Db 13469 GAGCCCTCTCGGCGCAGCAAGGCCCTCTGACTTTCAGGGATAGACTCTAGCAGCTTTAGC 13528
QY 2442 CCCAAGTAACTGAAGTCCACAGACACAGCAGAGATGTGGGGTCCAAAGGGCGGCCAC 2501
Db 13529 CCCAAGTAACTGAAGTCCACAGACACAGCAGAGATGTGGGGTCCAAAGGGCGGCCAC 13588
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGACCCGAGATAAGAC 2561
Db 13589 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCCTGACCCGAGATAAGAC 13648
QY 2562 AAAAGACCCGAGACAAATTTGAAACCTTCTCCAGTAGTCTCTTCCAGCCACCCCTCGG 2621
Db 13649 AAAAGACCCGAGACAAATTTGAAACCTTCTCCAGTAGTCTCTTCCAGCCACCCCTCGG 13708
QY 2622 CAGCGTAACATCAATGTTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 2681
Db 13709 CAGCGTAACATCAATGTTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 13768
QY 2682 TCCGGGAAGAGACTATTTCTGTAAATCGAGTGGCAGCAATCTGACAGCAATTTGGTGA 2741
Db 13769 TCCGGGAAGAGACTATTTCTGTAAATCGAGTGGCAGCAATCTGACAGCAATTTGGTGA 13828
QY 2742 GCCCTTCCAAAAAGACTGAAGTCAGCGTGGTGGCTTGAACCGTTGACCGCCGGGGC 2801
Db 13829 GCCCTTCCAAAAAGACTGAAGTCAGCGTGGTGGCTTGAACCGTTGACCGCCGGGGC 13888
```

```
QY 2802 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACT 2861
Db 13889 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACT 13948
QY 2862 GTTACCGCAGGACTGTGTGTATCCGTCGAGGCGTGTCTCCAAACCAAGGTTCTGTAG 2921
Db 13949 GTTACCGCAGGACTGTGTGTATCCGTCGAGGCGTGTCTCCAAACCAAGGTTCTGTAG 14008
QY 2922 CTCAGCAGAGGTGATTTCTCCAAATGTCTGACTTTTCCAGAGCCCTATGTTGGTCCGG 2981
Db 14009 CTCAGCAGAGGTGATTTCTCCAAATGTCTGACTTTTCCAGAGCCCTATGTTGGTCCGG 14068
QY 2982 GCCACTTTACACTTGTGTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 3041
Db 14069 GCCACTTTACACTTGTGTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAGAGTCT 14128
QY 3042 TGGTGGATGTGAGT 3055
Db 14129 TGCATGAGGGCGT 14142

RESULT 9
LOCUS AR070326 1507 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5892010.
ACCESSION AR070326
VERSION AR070326.1 GI:7221214
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray, J., Collins, C., Hwang, S.-i., Godfrey, T., Kowbel, D. and Rommens, J.
TITLE Genes from the 20Q13 amplicon and their uses
JOURNAL Patent: US 5892010-A 3 06-APR-1999;
FEATURES
source
1..1507
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 34.7%; Score 1106.2; DB 6; Length 1507;
Best Local Similarity 96.6%; Pred. No. 1.9e-299;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;
QY 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTCATGATGGATGG 60
Db 320 ATGCATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTCATGATGGAT-GG 378
QY 61 CCAGAGTGATTTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGCCCTTTGTCAATG 120
Db 379 CCAGAGTGATTTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGGAGATGGATG 437
QY 121 AAAGGGACCGCTGTTTTCATTTCCGAGCTACACAGAAATAATGTC-ATCCAAATCGA 179
Db 438 AAAGGGACCGCTGTTTTCATTTCCGAGCTACACAGAAATAATGTC-ATCCAAATCGA 497
QY 180 GGGGTATATGCCCTTTGGATTCGATTTCTGAGCCAGACCTTTCACATTTTCAGAGACCT 239
Db 498 GGGGAATATGCCCTTTGGATTCGATTTCTGAGCCAGACCTTTCACATTTTCAGAGACCT 557
QY 240 TAATAAATCATGCTTTAATGCAACCCGCCCTTACCTCTGTGAACCCAGCAGTCTTTCGGGT 299
Db 558 TAATAAATCATGCTTTAATGCAACCCGCCCTTACCTCTGTGAACCCAGCAGTCTTTCGGGT 617
QY 300 TGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGAACAGAACCTTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCCGCTTGTATAAAGTCAAGTGCAGAACAGAACCTTCCCAAGGA 677
QY 360 AAAGCAATTTGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGACATTTAGAGTCGC 419
Db 678 AAAGCAATTTGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGACATTTAGAGTCGC 737
```

```
Qy 420 TTTTGTGCTTTCAGATCCATGAGAACACACAAAGATTCTTTTCACTTACGGGTGTAACAT 479
Db 738 TTTTGTGCTTTCAGATCCATGAGAACACACAAAGATTCTTTTCACTTACGGGTGTAACAT 797
Qy 480 GTGCGGAAGAAGMTTSSRSSAGCCTTGTTCTTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAAGAAGATTCAAGGAGCCTTGTTCTTTTAAATAATCACATGCGGACRCRATAATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGACAGAGGCTTGGAGAGTAGTCAGCAACGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGACAGAGGCTTGGAGAGTAGTCAGCAACGATCAA 917
Qy 600 CGAGTCTGTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTACAAAATCTCATGGT 659
Db 918 CGAGTCTGTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTGCAAAATCTCATGGT 977
Qy 660 TTGTGGCTTCCCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAA 719
Db 978 TTGTGGCTTCCCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAA 1037
Qy 720 AAAAACTGCTTTCGGTACAGCGCGGCGACAGACTCTCCACAAGGAGGAATGCCGTC 779
Db 1038 AAAAACTGCTTTCGGTACAGCGCGGCGACAGACTCTCCACAAGGAGGAATGCCGTC 1097
Qy 780 CTCGAGGAGGACTTCTCTGCAAGTTGTTTCAACTTGAGACCAAAATCTCACCTGAAACGGG 839
Db 1098 CTCGAGGAGGACTTCTCTGCAAGTTGTTTCAACTTGAGACCAAAATCTCACCTGAAACGGG 1157
Qy 840 GAAGAAGCCTGTGAGATGATCCCTCAGCTCGATCGTTCACCACTTCCAGGCTTGCA 899
Db 1158 GAAGAAGCCTGTGAGATGATCCCTCAGCTCGATCGTTCACCACTTCCAGGCTTGCA 1217
Qy 900 GCTGCTACCAAGAAAGTTGCCATTTGCGAAGAGTGAAGAAATCGGGCGCAAGAGG 959
Db 1218 KCTGCTACCAAGAAAGTTGCCATTTGCGAAGAGTGAAGAAATCGGGCGCAAGAGG 1277
Qy 960 GAGCACGACAAACGACGATTCGAGTTCCGGAAGAGGCTTCGAGAAACAAATAAGGGCAG 1019
Db 1278 GAGCACGACAAACGACGATTCGAGTTCCGGAAGAGGCTTCGAGAAACAAATAAGAACCA 1337
Qy 1020 TTGTGAGGCTCTCGCAAGAAAGAGAGTGCACACACTCCACGCGCAAGCGCCCTC 1079
Db 1338 TTGTGAGGCTCTCGCAAGAAAGAGAGTGCACACACTCCACGCGCAAGCGCCCTC 1397
Qy 1080 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGGAGGCCACTCACTCTCGAGTG 1139
Db 1398 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGGAGGCCACTCACTCTCGAGTG 1457
Qy 1140 CGGCAAGCTTTCAGAACTTACCACGCTGGTCTTGCACCTCCAGGGTCC 1189
Db 1458 CGGCAAGCTTTCAGAACTTACCACGCTGGTCTTGCACCTCCAGGGTCC 1507
```

```
RESULT 10
BD085727
LOCUS BD085727 1507 bp DNA linear PAT 27-AUG-2002
DEFINITION Genes from the 20q13 amplicon and their uses.
ACCESSION BD085727
VERSION BD085727.1 GI:22631337
KEYWORDS JP 2001524802-A/3.
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 1507)
AUTHORS Gray,J.W., Collins,C.C., Hwang,S.I., Godfrey,T., Kowbel,D. and Rommens,J.
```

```
TITLE Genes from the 20q13 amplicon and their uses
JOURNAL Patent: JP 2001524802-A 3 04-DEC-2001:
THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
OS Artificial Sequence
PN JP 2001524802-A/3
PD 04-DEC-2001
```

```
PF 15-JUL-1997 JP 1998506264
PR 15-JUL-1996 US 08/680395,16-OCT-1996 US 08/731499 PR
17-JAN-1997 US 08/785532
PI JOE W GRAY, COLIN CONRAD COLLINS, SOO IN HWANG, TONY GODFREY, PI
DAVID KOWBEL,
PI JOHANNA ROMMENS
PC C12N15/11,C12Q1/68,A61K48/00
CC Description of Artificial Sequence:cc49
CC N is A, G, G, T, or U
FH Key modified base (447).
FT Location/Qualifiers
source
1..1507
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
ORIGIN
Query Match 34.7%; Score 1106.2; DB 6; Length 1507;
Best Local Similarity 96.6%; Pred. No. 1.9e-299;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;
Qy 1 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 60
Db 320 ATGCAATCGAAAGTGCAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGG 378
Qy 61 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGCGCTTGTCAATG 120
Db 379 CCAGAAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGA-GATGCTTGTCAATG 437
Qy 121 AAAGGAGCCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTG-ATCCAAATCGA 179
Db 438 AAAGGAGCCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCAATCCGAATCGA 497
Qy 180 GGGGTATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 239
Db 498 GGGGAATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 557
Qy 240 TAATAAACATGCTTAATGCAACACCGCCCTTACCTCTGTGAACAGCAGTTCCTCGGGT 299
Db 558 TAATAAACATGCTTAATGCAACACCGCCCTTACCTCTGTGAACAGCAGTTCCTCGGGT 617
Qy 300 TGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAAGTGCAGAACAGAACTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAAGTGCAGAACAGAACTCCCAAGGA 677
Qy 360 AAAGAAATGCAAGAAATGAATTTAGCTGTGAGTATGCGGAGAGATTTAGATGCGC 419
Db 678 AAAGAAATGCAAGAAATGAATTTAGCTGTGAGTATGCGGAGAGATTTAGATGCGC 737
Qy 420 TTTTGTGTTGAGATCCACATGAGAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 479
Db 738 TTTTGTGTTGAGATCCACATGAGAACACAAAGATTCTTTTCACTTACGGGTGTAACAT 797
Qy 480 GTGCGGAAGAAGMTTSSRSSAGCCTTGTTCTTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAAGAAGATTCAAGGAGCCTTGTTCTTTTAAATAATCACATGCGGACRCATAATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGACAGAGGCTTGGAGAGTAGTCAGCAACGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGACAGAGGCTTGGAGAGTAGTCAGCAACGATCAA 917
Qy 600 CGAGTCTGTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTCAAAAATCTTCATGGT 659
Db 918 CGAGTCTGTCAGGTGACCGCGCGGAGAGCATCTCTCTCTTTCGCAAAAATCTTCATGGT 977
Qy 660 TTGTGGCTTCCCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAA 719
Db 978 TTGTGGCTTCCCTATTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAA 1037
Qy 720 AAAAACTGCTTTCGGTACAGCGCGGCGACAGACTCTCCACAAGGAGGAATGCCGTC 779
Db 1038 AAAAACTGCTTTCGGTACAGCGCGGCGACAGACTCTCCACAAGGAGGAATGCCGTC 1097
```


Query Match	22.7%;	Score 722.6;	DB 2;	Length 208979;
Best Local Similarity	70.2%;	Prod. No. 5.6e-191;		
Matches 1106;	Conservative	0;	Mismatches 414;	Indels 55; Gaps 8;
Qy	1482	AGGTGAAAAACCATACAATGTGAATTTGTGAATATGTCAGACCCAGAACACATCTCT	1541	
Db	9324	AGGTGAAAAACCATACAATGTGAATTTGTGAGTATGCCGACGCCAGAACACATCTCT	99383	
Qy	1542	GAGGTATCACTTGGAGAGACATCAACAGGAAAAACAAAC--CGATGTTGCTGCTGAAGT	1598	
Db	99384	GAGGTACCACTTGGAGAGACATCAACAGACAGCAGCCGGTGATGTCGCCGCTGAGTC	99443	
Qy	1599	CAAGAACGATGGTAAAAATCAGGACACCTGAAGATGCATATTAA---CCGCTGACAGTGC	1655	
Db	99444	CAAAAGTGAAGCCGGAGCCAGGAGCCGAGGATGCGCTACTAAAGGCTGCTGACAGTGC	99503	
Qy	1656	GCAACCAAAATTTGAAAGATTTTTCATGGTGCCAAAGATGTTACAGGCAGTCCAC	1715	
Db	99504	GCAGACCAAAATTTAAAGAGATTTCTTGATGGTGCCAAAGATGTTAAGGGAAGGCCACC	99563	
Qy	1716	TGCAAGCAGCTTAAAGGAGATGCCCTCTGTGTTTTTCAGAAATGTTCTGGGCAGCGTGTCT	1775	
Db	99564	TGCCAAGCAGCTTAAAGGAGATGCCCTCTGTCTTCAGAGTGTC-----TCTC	99611	
Qy	1776	CTCACGACACACAAGATACTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGA	1835	
Db	99612	ACCAGCACACGAACGATACTCAGGATTTCCATAAACATGCAGCTGAT--AGTGCTGA	99668	
Qy	1836	TAAAGTGAATAAAAACCTTACCCCTGTCTTACCTGGACCTGTTAAAAAGAGATCAGCAGT	1895	
Db	99669	GAAGAAGGAGGAGAGCCCTGCCCTTACTTATCTGGACATGCAGAGAAAGAAA---GCAGG	99725	
Qy	1896	TGAAACTCAGGCAATAACCTCATCTGTAGAACCAAGCGCGAGTGTACTCTCTCTCCGA	1955	
Db	99726	GGAGCTCAGGCGACGACCCCTGTCTGCAGACTAGAGGGGGTGGGTCTTTAGCAGGGGA	99785	
Qy	1956	TGGCAGTACCACCCATAACCTTGAAAGTTAGCCCCAAAGAGAGAAACGAGACCGCAGC	2015	
Db	99786	AGCTGG-----CCATAGGAGAGAGATGATCAGGATGC	99818	
Qy	2016	TGACTGCAGATACAGGCCAAGTGTGGATTTGCAGAAACAACTTTAAATTTATCCGTGGG	2075	
Db	99819	TGACTACAGACATAAGCCCGGTGCTGACTGCCAGGACAGGCCCTTTGAAATCTATCCCTGG	99878	
Qy	2076	GGCTCTTACAAATTCGCCCGCAATTTCTTTGAGTAAAGTTTGATTCCAAGTATCACCTG	2135	
Db	99879	GCCGCTCACGCGCTGTCTCGCAATCTCTTTGAGCAAGTGTCTGATCCCCCAGCAITTCCTCG	99938	
Qy	2136	TCCATTTTGTAACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGCACAGAGACTGGA	2195	
Db	99939	CCCCCTTTGTACTTTCAAGACCTTTTATCCGGAAGTCTTATGATGCACAGAGACTTGA	99998	
Qy	2196	GCATAAATACAAATCTGACGTTTCATAAAACTGTCGAAACCAAGTCTTGTCTTAGAAGTCG	2255	
Db	99999	GCACAGGTACAACCTCGACCCGCAACAGAACCGGACGACGAAGTCTGTGCTCAGGAAACAG	100058	
Qy	2256	ACGTACCGGATGCCCGCAGCGGTGCTGGGAAAAAGATGTGCCCTCCCTCTCTAGTTTCTG	2315	
Db	100059	CGGTACCGGGGTGCCCTCCGGCTTTGCTGGGAAAAAGATGTCCCTCCCTTGTCTGGGCTGCA	100118	
Qy	2316	TAAACCAAGCCCAAGTCTGTCTTCCCGCGCAGTCCAAATCCCTCGCCATCTGCCGAAGG	2375	
Db	100119	CAAGCCCAAGGCCAAGATGCGCTTCTCACCACTCGAAGTCCCTTGCACTCAGAGAGGC	100178	
Qy	2376	GAAGCAGAGCCCTCCTGGGCCAGGCAAGGCCCTCTGTACTTCAGGGATAGACTCTAGCAC	2435	
Db	100179	TCGCGAGGGGGCTTCGGGGCCAAAGCAAGCAACCCAGACTTCAGGACCAGACAACAGCAC	100238	
Qy	2436	TTTAGCCCCAAGTAACCTGAAGTCCACAGACCAAGCAAGATGTGGGGTCCAGGGGCG	2495	
Db	100239	TTTAGCCCCAAGTAACCTGAAGTCAACAGGTGCAACCCCAATGTGGGGGCAACAG---100295		

QY	2496	CGCACACAGGCAACGAGCAATCTGAGATGTTTCTCTAAACCAAGTGTTCCTCCCTGCACCGGA	2555
Db	100296	TGCCACACAGGAGCAGCAGTCTGAGTGTCTTCCCAAGGTGGCG-TCTCTGCTGCTATGGA	100354
QY	2556	TAAGACAAAAGACCCGAGACAAAATTCGAAACCTCTTCCAGTAGTCTCTCTCTCAGGCCAC	2615
Db	100355	TAGGTGAAGAGACTGAGCCAAAACCTGAAGTCCCTACAGGCTCCCGCTCTCAGTCCCG	100414
QY	2616	CCTCGGACGAGTAACTCAATCTGTTTCCATCGACTACCCCGCAAGAACGACAGCCCGTG	2675
Db	100415	CCTCAGCAGTAATAATACCAAGGTTCTGTTGAGTATCCCGTGAAGTTGACGGCCCATG	100474
QY	2676	GGCACCTCCGGGAAGAGACTATTTCTGTAATCGGAGTGCCAGCAATCTCTCAGCAGAAAT	2735
Db	100475	GGCACAGCAAGGAGAGACTACTCTGCCATCGGAATCTGGCAGTCCCGCAGCAGAGTA	100534
QY	2736	TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTGTGCTTGTGAGTGTGACAGCC	2795
Db	100535	CAGTGAGCCATCCAAAAGACTCAAGTCCAGTCCGCTGCTCCCTGGACACAGAGCATGC	100594
QY	2796	CGGGGCCAATTACAGAGAGGCTATGACTCTTCCCAAGTACCATATGGTCAGAGGCATCAC	2855
Db	100595	AGGGACCAATGGCAGAGGGGCTTTGAGCTCCCAAGTACCATATGGTCAGAGGCATCAC	100654
QY	2856	ATCACTGTTACCGCAGGACTCTGTGTATCCGTTCGCGAGGCTGCCTCCCAACCAAGGTT	2915
Db	100655	CTCCTTGCTACCAACAGAGTGTGTGCGGCCACCGCTGTGCTGCCCAACCAAGCCGTTT	100714
QY	2916	CCTGAGCTCCACGAGGTCGATTCTCCAAATGTGTCTGACTGTTTCAGAGCCCTATGGTGG	2975
Db	100715	CCTGAGCCCTGGGAGTGGAGTACCCAGTGTGTGCTGTGTGAGAGCCCTCAGTGTC	100774
QY	2976	CTCGGGCCACTTTACACTTGTGTGCTGTGCTGTGAGTCCAGCATCCAGTCTCAGCGTTAGA	3035
Db	100775	CTCTGGACCCCTGTATACCTGTGAGCCCGTGGACACGACGAGGAGGCCAGCCCTTGA	100834
QY	3036	AGGTCTTGGTGGATG	3050
Db	100835	AGGTATTACTTAGG	100849
RESULT 13			
AC118373			
LOCUS			
DEFINITION	Rattus norvegicus clone CH230-356E19, *** SEQUENCING IN PROGRESS	linear	HTG 19-SEP-2002
ACCESSION			
VERSION	AC118373.3		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;		

Qy		2396	GGAATGATTTTCTGGAGTTCATAAATCCTCGAACAAATCTCTCCCTACCAATCTCTGCCTGTGCTGAGAGTCG	2235
Db		9999	GCACAGGTACAACCCTGACCGCACAGAACGGCAGCAGCAAGTCTGTGCTCAGGAACAG	100058
Qy		2256	ACGTAACGGGATTGCCCGCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTTTCTG	2315
Db		100059	CGGTACCGGGTGCCCTCCGGCTTTGCTGGGAAAAAGATGTCCCTCCCTTGTCTGGCGCTGCA	100118
Qy		2316	TAAACCCAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTTGCATCTGCCAAGGG	2375
Db		100119	CAAGCCCAAGGCCAAGACTGCTCTTCTACCAACACTCGAAGTCCCTGCATCTCAGAGAGGC	100178
Qy		2376	GAAGCAGAGCCCTCTGTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCNC	2435
Db		100179	TCGGCAGGGGGCTCTGGGGCCAAAGCAAAGCAGCCCCAGACTTCAGAGCCAGACAACAGCAC	100238
Qy		2436	TTTAGCCCCAAGTAACCTGAAGTCCCACAGACCACAGCAGAAATGTGGGGTCCAGGGGCG	2495
Db		100239	TTTAGCCCCAAGTAACCTGAAGTCCACAGGTTCACAAACCAGTGTGGGGGACACAG---	100295

REFERENCE
AUTHORS

1. (bases 1 to 184420)
Muzny, D., Marie, E., Metzker, M., Lee, S., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguitano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Blawlo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabis, A., Ganta, R., Garcia, A., Garner, F., Garza, M., Gebrgeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hoques,M., Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpachy,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshew,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A., Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E., Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Munitasa,M., Murphy,M., Nair,L., Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Paakeleneh,O., Okwuonu,G., Olarnpusagoon,A., Pal,S., Parks,K., Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rivers,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Vallas,R., Vera,V., Villanasa,D., Waidron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,R., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished
2 (bases 1 to 184420)
Worley,K.C.

Direct Submission

Submitted (17-APR-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 184420)

Rat Genome Sequencing Consortium.

Direct Submission

Submitted (19-SEP-2002)

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 19, 2002 this sequence version replaced gi:21745781.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GTWO

Center clone name: CH230-356E19

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 168728 bases at least Q40

Consensus quality: 170551 bases at least Q30

Consensus quality: 171717 bases at least Q20

Estimated insert size: 186818; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 20298: contig of 20298 bp in length
* 20299: gap of unknown length
* 20399: contig of 144604 bp in length
* 165002: gap of unknown length
* 165102: contig of 1811 bp in length
* 165103: gap of unknown length
* 166914: contig of 2456 bp in length
* 167014: gap of unknown length
* 169470: contig of 1645 bp in length
* 171214: gap of unknown length
* 171215: contig of 13106 bp in length.
* 171315: Location/Qualifiers

FEATURES

source

1. 184420
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-356E19"
1. 11115
/note="wgs_contig"

misc_feature

ORIGIN

Query Match	22.5%;	Score 716.2;	DB 2;	Length 184420;
Best Local Similarity	69.7%;	Pred. No. 3.4e-119;	Indels 60;	Gaps 7;
Matches 1097;	Conservative 0;	Mismatches 418;		

QY 1482 AGGTGAAAACCATACATAAATGTGAATTTTGTGAATATGTCTCAGCCCAAGAGACATCTCT 1541

Db 146201 AGGTGAAAACCATACATAAATGTGAATTTTGTGAATATGTCTCAGCCCAAGAGACATCTCT 146260

QY 1542 GAGGTATCCTTTGGAGAGACATCAAGAGAAAACA---AACCGATGTTGCTGCTGAAGT 1598

Db 146261 GAGGTATCCTTTGGAGAGACATCAAGAGAAAACA---AACCGATGTTGCTGCTGAAGT 146320

QY 1599 CAAGACGATGTTAAAATCAGGACATCAAGAGAAAACA---AACCGATGTTGCTGCTGAAGT 1658

Db 146321 CAAGACGATGTTAAAATCAGGACATCAAGAGAAAACA---AACCGATGTTGCTGCTGAAGT 146380

QY 1659 AACCAAAAATTTGAAAAGATTTTGTGATGTTGTCGCAAGATGTTACAGGCAGTCCACCTGC 1718

Db 146381 GACCAAAAATTTAAGAGATTTCTTGATGTTGTCGCAAGATGTTAAGGCGAGCCCAACCCGC 146440

QY 1719 AAAGCAGTTAAGAGATGCTCTCTGTTTTCAGATGTTCTGGGCGAGCGTCTCTCTC 1778

Db 146441 CAAGCTACTTAAGAGATGCTCTCTGTTTTCAGATGTTCTGGGCGAGCGTCTCTCTC 146500

QY 1779 ACCAGCAC---ACAAAGATCTCAGGATTTCCATAAAAATCAGCTGATGACAGTCTGA 1835

Db 146501 ACCAGCACAAACAGGATCTCAGGATTTCCATAAAAATCAGCTGATGACAGTCTGA 146557

QY 1836 TAAAGTGAATAAAAACCTTACCTGCTTACCTGACCTGTTTAAAAAAGATCAGCAGT 1895

Db 146558 GAAAGTGCAGAGAGTCTGCGCCCTGCTTATCTGACATGCCAAGAAAGA---AAACAGG 146614

QY 1896 TGAAGCTCAGCAAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCCGA 1955

Db 146615 GGATCCCGGGCGAGCAGCCCTGCTGCAAACTAGAGGGGGTGGGCCCTTAGCACAGAGA 146674

QY 1956 TGCGAGTACACCCCAATACCTTGAAGTTAGCCCAAGAGCAAAACGGAGACCGCAGC 2015

Db 146675 AGCTGG-----CCACGGGAGAGATGGATCGGAGGC 146707

QY 2016 TGACTGCAGATACAGGCCCAAGTGTGGATTGTCAAAAAAACCCTTTAAATTTATTCGGTGG 2075

Db 146708 TGACTGCAGATACAGGCCCGGTGCTGACTGCCAGCAGAGCCCTTTGAATCTATCACTCGG 146767

QY 2076 GGCTCTTCAAAATGGCCGGCAATTTCTTTGAGTAAAGATTTGATTCAGATATCACCTG 2135

Db 146768 GGCACCTCCAGCGCTGCTCCTGCAATCTCTTTGAGCAAGTGTCTCATCCCTAGCATTTGCCTG 146827
QY 2136 TCAATTTTGTACTCTCAAGACATATTTATCCAGAAGTTTAAATGATGACGACGAGACTGA 2195
Db 146828 CCAATTTTGTACTCTTTAAACCTTTTATCCGAAGTCTCTATGATGACGACGAGACTTGA 146887
QY 2196 GCATAAATAACATCTCGAGCTTCATAAATACTGCGAACAAGTCTTGTCTTAGAGTCG 2255
Db 146889 GCACAGGTACAACCCCGACGACGACAGAAGACTGCGAGCAAGTCTGTGCTGAGGAACG 146947
QY 2256 AGCTACCGGATCCCGCCAGCGTGTCTGGAAAAAGATGTCCTCCCTCTCTAGTTTCTG 2315
Db 146948 CGGACAGGCTGCCCTCTGCTGCTGGGAAGATGATACCCCTCTGTCTGCTGCTGA 147007
QY 2316 TAAACCAAGCCCAAGTCTGCTTTCCCGGCGGAGTCCAAATCCCTGCCATCTGCGAAGG 2375
Db 147008 CAAGCCCAAGGCGCAAGCTGCTCTCATCAGAGGCTAAGGCGCTGCACCCAGAGAAGT 147067
QY 2376 GAAGCAGAGCCCTCTCGGCGCAGGCAAGGCCCTCTGACTTCAGGATAGACTCTAGCAC 2435
Db 147068 TCGGACAGGGAACCTCGGGCCAGGCAAGCTCCCGAGACTTCAGGACCGGACAACAGCAC 147127
QY 2436 TTTAGCCCCAAGTAACCTGAAGTCCACAGACACAGCAGAGATGTGGGGGTCCAAAGGGGC 2495
Db 147128 TTTAGCCCCAAGTAACCTGAAGTCCACAGAGTCCAGGCCAGCTGCGGGGCCCCAG--- 147184
QY 2496 CGCCACAGGCAACAGCAATCTGAGATGTTTCTTAAACAGTGTTCCTTCCCTGCACCGGA 2555
Db 147185 TGCCACAGGCGAGCAGCAGTCAAGCTGTTTCCAAAAGTGGCGTCCCTGCTGCTATGGA 147244
QY 2556 TAAAGCAAAAGACCCGAGACAAAATGAACTCTTCAGTAGTCTCTTCTTCAGCCAC 2615
Db 147245 TAAAGCGAAAGACCTGAGCCCAAACTGAAGCCCGCCCGGCTCCCGGCTCAGTCCCC 147304
QY 2616 CCTCGCAGCAGTAACTCAATGGTTCCATGCTACTACCCGCAAGAACGACGAGCCGCTG 2675
Db 147305 CTTAGCAGTGTATGCGAAGGTTCT-----GGCTAGTGCCTATG 147346
QY 2676 GGCACCTCCGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGCAGCAGAATT 2735
Db 147347 GGGGAGCAAGGAAGAGACTACTGTCTCAACGGAATGCTGGCAGTGCACGCTCAGNATA 147406
QY 2736 TGGTGAAGCCCTCCAAAAGACTGAAGTCCAGCGTGTGGTCCCTTGCAGTTGACCGCC 2795
Db 147407 CAGTGAGGCACATCCCAAGACTCAAGTCCCTGTGGTGTCTCCCTGGACACAGAGCATCC 147466
QY 2796 CGGGGCCAAATACAGAAGAGGCTATGACCTTCCCAAGTACCATATGTTCCAGAGGCATCAC 2855
Db 147467 TGGGCCCAATGCGAGAGGGGCTTTGAGCTCCCAAGTACCATGTTGTTGAGGAGCATCAC 147526
QY 2856 ATCACTGTTTACCGCAGGACTGTGTATCCGTGCGAGGGCGTCCCTCCCAACCCAGGTT 2915
Db 147527 TTCCCTGTTTACCACAGAGTATGTGTCTCACCGCTGTGCTGCTCCCAAGGCCGTTT 147586
QY 2916 CTTGAGCTCCAGGAGTGCATTTCTCAATGTGCTGACTGTTTCAGAGCCCTATGGTGG 2975
Db 147587 CTTGAGCCCTGGGAGGTGGAGTCCGCCAGCGTGTGACTGTGCGAGAAGCCCTACGGTGC 147646
QY 2976 CTTCCGGGCCACTTTACATTTGTCCTGCTGCTGCTAGTCCAGCATCCAGCTCGACGTTAGA 3035
Db 147647 CTTGAGCCCTGTATACCTGTGGACCCGTTGGACATGCGAGGACCGCCATCCCTTGA 147706
QY 3036 AGGTTCTTGTGGATG 3050
Db 147707 AGGTATTACTGTAGG 147721

RESULT 14

AC099079 246332 bp DNA linear HTG 10-MAY-2003
LOCUS Rattus norvegicus clone CH230-99J9, *** SEQUENCING IN PROGRESS ***
DEFINITION 6 unordered pieces.
ACCESSION AC099079

VERSION
KEYWORDS
SOURCE
ORGANISM

AC099079.5 GI:30521728
HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
Rattus norvegicus (Norway rat)
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS

1 (bases 1 to 246332)
Munzy,D.Marie., Metzker,M.Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleaveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gumaratne,P., Healand,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jollivet,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Johnson,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Liu,J.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenschwari,L., Louissegh,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapa,P., Martin,K., Martin,R., Martinez,E.,
Munhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwackemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu.L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,R., Rose,R., Ruiz,S.J.,
Sanders,M., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorrelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Uemami,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu.F., Zhang,J., Zhou,X., Zhou,X., Zhao,S., Dunn,D., von
Niederhauser,A., Weise,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 246332)
Worley,K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 246332)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23269227.
The sequence in this assembly is a combination of BAC based reads

REFERENCE
AUTHORS
TITLE
JOURNAL

3 (bases 1 to 246332)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23269227.
The sequence in this assembly is a combination of BAC based reads

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: GF2C
Center clone name: CH230-99J9
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 209040 bases at least Q40
Consensus quality: 21312 bases at least Q30
Consensus quality: 216534 bases at least Q20
Estimated insert size: 219845; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_draft\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 14481: contig of 14481 bp in length
* 14482: gap of unknown length
* 14582: contig of 26430 bp in length
* 41012: gap of unknown length
* 41112: contig of 198708 bp in length
* 239820: gap of unknown length
* 239920: contig of 1203 bp in length
* 241123: gap of unknown length
* 241223: contig of 1094 bp in length
* 242317: gap of unknown length
* 242417: contig of 3916 bp in length.
Location/Qualifiers
1..246332
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-99J9"
/misc_feature
1..4672
/note="wgs_contig"
misc_feature
136991..139129
/note="wgs_contig"
misc_feature
174114..175268
/note="wgs_contig"

ORIGIN
Query Match 22.5%; Score 716.2; DB 2; Length 246332;
Best Local Similarity 69.7%; Pred. No. 3.7e-189;
Matches 1097; Conservative 0; Mismatches 418; Indels 60; Gaps 7;
QY 1482 AGTGAAACCATCAAAATGTAATTTGTAATATGTCGAGCCCAAGAGACATCTCT 1541
DB |||||
DB 55734 AGTGAAACCATCAAAATGTAATTTGTAATATGTCGAGCCCAAGAGACATCTCT 55793
QY 1542 GAGGTATCATTGGAGAGACATCAAGGAAACAA---ACCAGATTTGCTGCTGAAGT 1598
DB |||||
DB 55794 GAGGTATCATTGGAGAGACATCAAGGAAACAA---ACCAGATTTGCTGCTGAAGT 55853
```

```
QY 1599 CAAGAACGATGGTAAATAATCAGGACACTGAAGATGCACTATTAAACCCCTGACACTGGCGCA 1658
DB |||||
DB 55854 CAAAGTGAAGCCGCGAAGTCAGGAGCCGCAAGATATGCTACTAACGGCTGATAGTGACA 55913
QY 1659 AACCAAAATTTGAAAGATTTTTTGATGTGCAAAAGATGTTACAGGCAGTCACCTGC 1718
DB |||||
DB 55914 GACCAAAATTTAAAGAGATTTCTTTGATGTGCAAAAGATGTTAAAGGCGACGCCACCGC 55973
QY 1719 AAAGCAGCTTAAGAGAGATGCTCTCTGTTTTTCAAGATGTTCTGGGCGAGCGCTGCTCTC 1778
DB |||||
DB 55974 CAAGCTACTTAAGAGAGATGCTCTCTGTTTTTCCAGAGTGTCTGGGCGAGCACTGTTCTCTC 56033
QY 1779 ACCAGCAC---ACAAAGATATCTCAGGATTTTCCATAAAATATGACGCTGATGAGTGTGA 1835
DB |||||
DB 56034 ACCAGCACACAAACAGGATATCTCAGGATTTTCCATAAAACGCGAGCTGATG---GTGCTGA 56090
QY 1836 TAAAGTGAATAAAACCCCTACCTGCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGT 1895
DB |||||
DB 56091 GAAAGTGCAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56147
QY 1896 TGAAGCTCAGGCAATAAATACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCTCCGGA 1955
DB |||||
DB 56148 GGATCCCGGCGCAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56207
QY 1956 TGGCAGTACCCACCCATTAACCTTTGAAGTTAGCCCCCAAGAGAGAACCGGAGACCGCAGC 2015
DB |||||
DB 56208 AGCTGG-----CCACGGGAGAGATGGATCGGAGGC 56240
QY 2016 TGACTGCAGATACAGGCAAGTGTGAGATGTCAGCAAAAACCTTTAAATTTATCGTGGG 2075
DB |||||
DB 56241 TGACTACAGACATGAAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56300
QY 2076 GGCTCTTCACAATTCGCGGCAATTTCTTTGAGTAAAGTTTGAATTCAGATATCACTG 2135
DB |||||
DB 56301 GGCACTCAGCCCTGCTGCAATCTCTTTGAGCAGTGTCTCATCTAGTAGTGGCTG 56360
QY 2136 TCCATTTTGTACCTTCAAGACATTTATCCAGAAAGTTTATATGATGACCAAGAGACTGGA 2195
DB |||||
DB 56361 CCCATTTTGTACTTTTAAACCTTTTATCCGGAAGTCTTATGATGACCAAGAGACTGGA 56420
QY 2196 GCATAATACATCTCTGACGTTTCAATAAACTGCAAAAGTCTTGAAGTCTTGAAGTCTG 2255
DB |||||
DB 56421 GCACAGGTACAAACCCGACAGCACAAGAACTGAGCAGCAAGTCTGTGCTGAGAGAACAG 56480
QY 2256 ACGTACCGAGTGCCTGCGGCGAGGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTCTG 2315
DB |||||
DB 56481 CGGACAGGGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 56540
QY 2316 TAAACCCCAAGCCCAAGTCTGCTTTCCGCGGCGAGTCCCAATTCCTGCAATCTGCAAGGG 2375
DB |||||
DB 56541 CAAGCCCAAGCCCAAGCTGCTTCTCATCACAGGCTAAGGCCCTGACCCACAGAGAGAGT 56600
QY 2376 GAAGCAGAGCCCTCTCTGGGCGAGGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCAC 2435
DB |||||
DB 56601 TCGGCGAGGGACCTCTGGGGCCAGGCAAAAGTCTCCGAGACTTCAGGACCGGACAAACAGCAC 56660
QY 2436 TTTAGCCCCCAAGTAACTGAGTCCACAGACACAGAGATGTGGGGTCCCAAGGGGC 2495
DB |||||
DB 56661 TTTAGCCCCCAAGTAACTGAGTCCACAGAGTCAAGCCAGTGCAGGGGGGCCAG--- 56717
QY 2496 CGCCACCAAGGCAACAGCAATCTGAGATGTTTTTCTTAAACCAAGTGTTCCTCTGCAACCGGA 2555
DB |||||
DB 56718 TGCCACCAAGGCGAGCAGCAGTCAAGAGTGTTCCTTAAAGAGTGGCTGCTGCTGCTGCTGCTG 56777
QY 2556 TAAGCAAAAAGACCCGAGACAAAATTAACCTTCCAGTAGTCTCTCTCAGCCAC 2615
DB |||||
DB 56778 TAAGCGGAAAAGACCTGAGGCCAAACTGAAGGCCCGCCCGCTCCCCGGCTCAGTCCCC 56837
QY 2616 CCTCGGAGCAGTAACTCAATGTTCCATCGACTACCCCGCCCAAGACAGCAGCCGCTG 2675
DB |||||
DB 56838 CCTTAGCAGTGTATGCAACGGTTC-----GGCTGAGTGCCCATG 56879
```

```
QY 2676 GGCACCTCCGGGAAGAGACTATTCTGTAAATCGGAGTGCCAGCAATCTGTCAGCAGAAATT 2735
|||
Db 56880 GGGCAGCAGCAAGAGAGACTACTCTGTCAACGGATGTCGAGTGCCACGCTCAGAATA 56939
|||
QY 2736 TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGGTGGTGGCCCTTGAGCTTGACACGACC 2795
|||
Db 56940 CAGTGAGCCACATCCCAAGACTCAAGTCCCTGTGGTGTCCCTGGACACAGAGCATCC 56999
|||
QY 2796 CGGGGCCAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGCTCAGAGGCATCAC 2855
|||
Db 57000 TGGGGCCCAATGGCAGAGGGGCTTTGAGCTCCCAAGTACCATATGCTCAGAGGCATCAC 57059
|||
QY 2856 ATCACTGTTACCGCAGGACTGTGTATCCGTCCGAGGGCGCTGCCCTCCCAAAACCAAGGTT 2915
|||
Db 57060 TTCCCTGTTACCAACAGAGTATGTGCTCACGGCTGTGCTCTCCCAAGACCCGTTT 57119
|||
QY 2916 CTGAGCTCCAGCGAGTGCATTCCTCAAAATGTGTGACTGTTCAGAGCCCTATGGTGG 2975
|||
Db 57120 CCTGAGCCCTGGGAGGTGGAGTCCGCCACGCGTGTGACTGTGCAGAGCCCTACGGTGC 57179
|||
QY 2976 CTCGGGGCCACTTTACACTGTGTGCTGCTGCTAGTCCAGCATCCAGCTCGACGTTAGA 3035
|||
Db 57180 CTCTGAGCCCTGTATACCTGTGGACCGCGTGGACATGAGGAGCCGACATCCCTTTGA 57239
|||
QY 3036 AGGTCTTGGTGGATG 3050
|||
Db 57240 AGGTATTACTGTAGG 57254
|||
```

RESULT 15

```
BD060160
LOCUS BD060160 469 bp DNA linear PAT 27-AUG-2002
DEFINITION Secreted expressed sequence tags (seSTs).
ACCESSION BD060160
VERSION BD060160.1 GI:22605766
KEYWORDS JP 2001518793-A/520.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 469)
Treacy,M., Mccoy,J.M., Lavallie,E.R., Racie,L.A., Merberg,D.,
Treacy,M., Spaulding,V. and Agostino,M.J.
Secreted expressed sequence tags (seSTs)
Patent: JP 2001518793-A 520 16-OCT-2001;
GENETICS INSTITUTE INC
PN JP 2001518793-A/520
PD 16-OCT-2001
PF 16-APR-1998 JP 1998543070
PR 10-APR-1997 US 08/837312
PI KENNETH JACOBS,JOHN M MCCOY,EDWARD R LAVALLIE,LISA A RACIE, PI
DAVID MERBERG,
PI MAURICE TREACY,VIKKI SPAULDING,MICHAEL J AGOSTINO PC
C12N15/12,C12N5/10,C07K14/47,C12Q1/68,A61K38/17 CC Strandedness:
Double;
CC topology: Linear;
FH Key Location/Qualifiers.
FEATURES
source
1..469
/organism="Zea mays"
/mol_type="genomic DNA"
/db_xref="taxon:4577"
```

ORIGIN

```
Query Match 13.9%; Score 444.4; DB 6; Length 469;
Best Local Similarity 99.8%; Pred. No. 3.3e-113;
Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1032 CTCGCAAGAGAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCCGTGGACGCGGA 1091
|||
Db 19 CTAGCAAGAGAAGAGAGTGCAAACACTCCACGCGGAGCGCCCTCCGTGGACGCGGA 78
|||
```

```
QY 1092 TCCCAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGGCCAAAGCTTT 1151
|||
Db 79 TCCCAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGGCCAAAGCTTT 138
|||
QY 1152 CAGAACCTTACCAACAGCTGGTCTTGGCACTCCAGGGTCCCAAGAAAGGACCGGAGGCGCG 1211
|||
Db 139 CAGAACCTTACCAACAGCTGGTCTTGGCACTCCAGGGTCCCAAGAAAGGACCGGAGGCGCG 198
|||
QY 1212 CGCGAGTGGCCACCATGTCTGTGGACGGGAGGAGCCGGGGACGTGTTCTCTGACCT 1271
|||
Db 199 CGCGAGTGGCCACCATGTCTGTGGACGGGAGGAGCCGGGGACGTGTTCTCTGACCT 258
|||
QY 1272 CGCGCCCTCTTGGATGAAATGGAGCCGTGGATCGAGGGAGGCTTCTTGAAGACGG 1331
|||
Db 259 CGCGCCCTCTTGGATGAAATGGAGCCGTGGATCGAGGGAGGCTTCTTGAAGACGG 318
|||
QY 1332 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAGGAGGAAAAAT 1391
|||
Db 319 ATCTGAGGATGGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAGGAGGAAAAAT 378
|||
QY 1392 AAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGGTTTTTCCGTTCAAATTA 1451
|||
Db 379 AAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGGTTTTTCCGTTCAAATTA 438
|||
QY 1452 TTACCTCAATATTCAATCTCAGAACGC 1477
|||
Db 439 TTACCTCAATATTCAATCTCAGAACGC 464
|||
```

Search completed: July 4, 2005, 01:46:53

Job time : 13520.8 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 05:45:02 ; Search time 1627.82 Seconds

(Without alignments)
11586.223 Million cell updates/sec

Title: US-08-731-499-10

Perfect score: 3186

Sequence: 1 ATGCAATCGAAGTGACAGG.....GGACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3183.6	99.9	5252	11 ACN44987	Acn44987 Human mrn
2	3156.4	99.1	3183	2 AAV09024	AAV09024 Homo sapi
3	3040.6	95.4	5632	8 ACC50334	ACC50334 Breast ca
4	3040.6	95.4	5632	10 ADD14635	Add14635 Human src
5	3040.6	95.4	5632	13 ADR66761	Adr66761 Human pro
6	3040.6	95.4	5632	13 ADR65858	Adr65858 Human pro
7	3040.6	95.4	5653	12 ADP07267	Adp07267 Human ZNF
8	1561.2	49.0	14906	4 AAK81093	AAK81093 Human imm
9	1561.2	49.0	36022	11 ACN44986	Acn44986 Human gen
10	1511.2	47.4	3016	11 ACN44985	Acn44985 Mouse mrn
11	1482.8	46.5	10282	2 AAV09023	AAV09023 Homo sapi
12	1106.2	34.7	1507	2 AAV04698	AAV04698 Homo sapi
13	733	23.0	26345	11 ACN44984	Acn44984 Mouse gen
14	563.6	17.7	566	4 AAK81095	AAK81095 Human imm
15	444.4	13.9	469	2 AAV88042	AAV88042 EST clone
16	401.4	12.6	530	12 ACH90216	Ach90216 Human gen
17	251	7.9	251	12 ACH90216	Ach90216 Human gen
18	104	3.3	5994	13 ACN40355	Acn40355 Tumour-as
19	102.4	3.2	6033	13 ADR25650	Adr25650 Breast ca
20	102.4	3.2	8156	6 ABT07614	Abt07614 Human bre

C	21	72	2.3	424	3	AAA42306	Aaa42306 Human sec
	22	72	2.3	2706	10	ADP07836	Adp07836 Novel cod
	23	70	2.2	70	13	ADR00072	Adr00072 ZNF217 PC
C	24	66.8	2.1	1014	12	ACH87715	Ach87715 Human gen
C	25	66.8	2.1	1229	12	ACH86998	Ach86998 Human gen
	26	66.8	2.1	2706	11	ADL71050	Adl71050 Gene enco
	27	66.8	2.1	2765	4	AAI58185	Aai58185 Human pol
	28	66.8	2.1	2765	5	ADQ98391	Adq98391 DNA enco
	29	66.8	2.1	2765	9	ADB48151	Adb48151 Novel hum
	30	66.8	2.1	3039	13	ADR25783	Adr25783 Breast ca
	31	66.8	2.1	3040	13	ADQ87590	Adq87590 Human tum
	32	66.8	2.1	3617	4	ABK42507	Abk42507 Genomic s
	33	66.8	2.1	3617	4	ABK42509	Abk42509 Genomic s
	34	66.8	2.1	3617	9	ADB60663	Adb60663 Connectiv
	35	66.8	2.1	3617	9	ADB60665	Adb60665 Connectiv
	36	64.2	2.0	2909	11	ADL71048	Adl71048 Gene enco
	37	63.8	2.0	2109	10	ADJ80195	Adj80195 Novel hum
	38	62.8	2.0	1033	12	ACH89764	Ach89764 Human gen
	39	62.8	2.0	2223	5	AA89101	Aas89101 DNA enco
	40	62.8	2.0	2918	10	ADB62681	Adb62681 Human CDN
	41	62.8	2.0	4599	4	AAK90261	Aak90261 Human dig
	42	62.8	2.0	4599	4	AAK90264	Aak90264 Human dig
	43	62.8	2.0	4599	5	AA839893	Aas39893 Genomic s
	44	62.8	2.0	4599	5	AA839890	Aas39890 Genomic s
	45	62.8	2.0	4599	9	ADB32853	Adb32853 Human nov

ALIGNMENTS

RESULT 1

ACN44987

ID ACN44987 standard; cDNA; 5252 BP.

XX ACN44987;

XX ACN44987;

DT 18-NOV-2004 (first entry)

XX Human mRNA sequence hCT1950762.

DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ds.

XX Homo sapiens.

XX WO2003073826-A2.

FN 12-SEP-2003.

PD 28-FEB-2003; 2003WO-US006235.

XX 01-MAR-2002; 2002US-00087192.

XX (SAGR-) SAGRES DISCOVERY.

XX Morris DW;

XX WPI; 2003-328604/31.

XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma

XX PT comprises a nucleotide sequence.

XX Claim 1; SEQ ID NO 1709; Opp; English.

XX The present invention relates to novel DNA and protein sequences which

XX are associated with carcinomas. The sequences are useful for: (i) for

XX screening drug candidates; (ii) for screening of bioactive agent capable

XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of

XX a bioactive agent capable of modulating the activity of CAP; (iv) for

XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing

XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating

XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;

XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for

XX determining Carcinoma Associated (CA) gene copy number. In addition, the

CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA gene
CC sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published

SQ Sequence 5252 BP; 1507 A; 1101 C; 1177 G; 1467 T; 0 U; 0 Other;

Query Match 99.9%; Score 3183.6; DB 11; Length 5252;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy	1	ATGCAATCGAAGTGACAGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG	60
Db	25	ATGCAATCGAAGTGACAGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG	84
Qy	61	CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG	120
Db	85	CCAGAAGTGATTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCTTGTCAATG	144
Qy	121	AAAGGGACCGCTGTGTTCCATTCCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG	180
Db	145	AAAGGGACCGCTGTGTTCCATTCCGAGCTACACAAGAAAAAATGTCAATCCAAATCGAG	204
Qy	181	GGGTATATGCCCTTGGATGTGATGTTCTGACGCGAGACCTTCAACATTCAGAAACCTT	240
Db	205	GGGTATATGCCCTTGGATGTGATGTTCTGACGCGAGACCTTCAACATTCAGAAACCTT	264
Qy	241	AATAAATCATGTTTAAATCAACACCGGCTACCTCTGTAACGACGAGTCTTCGGGTT	300
Db	265	AATAAATCATGTTTAAATCAACACCGGCTACCTCTGTAACGACGAGTCTTCGGGTT	324
Qy	301	GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA	360
Db	325	GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA	384
Qy	361	AAGAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT	420
Db	385	AAGAATTCGAAGAAAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT	444
Qy	421	TTTGATGTTGAGATCCATGAGACACACAAAGATTTCTTCAATTCAGGCTGTAAACATG	480
Db	445	TTTGATGTTGAGATCCATGAGACACACAAAGATTTCTTCAATTCAGGCTGTAAACATG	504
Qy	481	TGCGAAGAAAGMTTSSRSAGCTTGGTCTTAAATAATCAATGCGGACACATAATGGC	540
Db	505	TGCGAAGAAAGMTTCAAGAGCTTGGTCTTAAATAATCAATGCGGACACATAATGGC	564
Qy	541	AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC	600
Db	565	AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC	624
Qy	601	GAGGTGCTCCAGGTGACGCGCGAGAGCATCTCTCTTACAAATCTGCATGGTT	660
Db	625	GAGGTGCTCCAGGTGACGCGCGCGAGAGCATCTCTCTTACAAATCTGCATGGTT	684
Qy	661	TGTGCTCTCTTATTTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAA	720
Db	685	TGTGCTCTCTTATTTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAA	744
Qy	721	AAAATGCTTTGCTGATGATGATCCCTCAGCTCGATCCGTTTCAACAGGAGGAATGCCGTC	780
Db	745	AAAATGCTTTGCTGATGATGATCCCTCAGCTCGATCCGTTTCAACAGGAGGAATGCCGTC	804
Qy	781	TCGAGGGAGGACTTCTCCAGTGTGTTCAATTTGAGACCAAAATCTCACCTTGAACCGGG	840
Db	805	TCGAGGGAGGACTTCTCCAGTGTGTTCAATTTGAGACCAAAATCTCACCTTGAACCGGG	864
Qy	841	AAGAAGCTGTGATGATGATCCCTCAGCTCGATCCGTTTCAACAGGAGGAATGCCGTC	900
Db	865	AAGAAGCTGTGATGATGATCCCTCAGCTCGATCCGTTTCAACAGGAGGAATGCCGTC	924
Qy	901	CTGGCTACCAAGGAAAAATGTCCTTCCAAAGTGAAGGAATCCGGGGCAAGAGGG	960

Db	925	CTGGCTACCAAGGAAAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGCAAGAAGGG	984
Qy	961	AGCACCCACACACGAGATTCGAGTTCCGAGAAGGAGCTTGGAGAAAACAAATAAGGGCAGT	1020
Db	985	AGCACCCACACACGAGATTCGAGTTCCGAGAAGGAGCTTGGAGAAAACAAATAAGGGCAGT	1044
Qy	1021	TGTGACGGCTCTCTCGCAAGAGAAAGAGAAGTGCNAACACATCCCAACGCGAAGCCCTTC	1080
Db	1045	TGTGACGGCTCTCTCGCAAGAGAAAGAGAAGTGCNAACACATCCCAACGCGAAGCCCTTC	1104
Qy	1081	GTGACGCGGATCCCAAGATTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGC	1140
Db	1105	GTGACGCGGATCCCAAGATTACCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTGC	1164
Qy	1141	GGCAAAAGCTTTCAAGAACCTACCAACAGCTGCTCTTGACCTCCAGGGTCCCAAGAGGAC	1200
Db	1165	GGCAAAAGCTTTCAAGAACCTACCAACAGCTGCTCTTGACCTCCAGGGTCCCAAGAGGAC	1224
Qy	1201	CGGAGGCGCGCGGAGTCCGCCACCATGTTCTGTGACGGGAGGACCGGGGACGTGT	1260
Db	1225	CGGAGGCGCGCGGAGTCCGCCACCATGTTCTGTGACGGGAGGACCGGGGACGTGT	1284
Qy	1261	TCTCTGACCTCGCGCGGCTCTGGATGAAAATGGAGCGCTGGATCCGAGGGGAAGTGGT	1320
Db	1285	TCTCTGACCTCGCGCGGCTCTGGATGAAAATGGAGCGCTGGATCCGAGGGGAAGTGGT	1344
Qy	1321	TCTGAACAGCGATCTGAGGATGGCTTCCGGAAGGAATCCATCTGGATTAATAATGATGAT	1380
Db	1345	TCTGAACAGCGATCTGAGGATGGCTTCCGGAAGGAATCCATCTGGATTAATAATGATGAT	1404
Qy	1381	GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTATTTCTGGAAAAGTTTTTC	1440
Db	1405	GGAGGAAAAATAAACAATCTTACATCTTCAAGAGAGTGTATTTCTGGAAAAGTTTTTC	1464
Qy	1441	CGTTCAAAATTTATACCTCAATATTTATCTCAGAACGATACAGGTGAAAACCATACAAA	1500
Db	1465	CGTTCAAAATTTATACCTCAATATTTATCTCAGAACGATACAGGTGAAAACCATACAAA	1524
Qy	1501	TGTCAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA	1560
Db	1525	TGTGAATTTTGTGAATATGCTGACGCCAGAGACATCTCTGAGGTATCACTTCGAGAGA	1584
Qy	1561	CATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGTTAAAAATCAG	1620
Db	1585	CATCACAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAACGATGTTAAAAATCAG	1644
Qy	1621	GACACTGAAGTGCATCTTAAACCGCTGACAGTGGCGAAAACCAAAATTTGAAAAGATTT	1680
Db	1645	GACACTGAAGTGCATCTTAAACCGCTGACAGTGGCGAAAACCAAAATTTGAAAAGATTT	1704
Qy	1681	TTTGATGTCGCAAGATGTTTACAGGACGTCACCTGCAAGACGAGCTTAAGGAGATGCCT	1740
Db	1705	TTTGATGTCGCAAGATGTTTACAGGACGTCACCTGCAAGACGAGCTTAAGGAGATGCCT	1764
Qy	1741	TCTGTTTTTCAAGATGTTTGGGCGAGCTGTCTCTCAACAGCACCAAAAGATCTCAG	1800
Db	1765	TCTGTTTTTCAAGATGTTTGGGCGAGCTGTCTCTCAACAGCACCAAAAGATCTCAG	1824
Qy	1801	GATTTCCATAAAAATGACGCTGATGACAGTGTCTGATAAAGTGAATAAAAACCTTACCCCT	1860
Db	1825	GATTTCCATAAAAATGACGCTGATGACAGTGTCTGATAAAGTGAATAAAAACCTTACCCCT	1884
Qy	1861	GCTTACCTGGAACCTGTTTAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAAATCTCATC	1920
Db	1885	GCTTACCTGGAACCTGTTTAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAAATCTCATC	1944
Qy	1921	TGTAGAACCAAGGCGGATGTTTACTCTCTCCCGGATGGCAGTACCAACCATTAACCTTGAA	1980
Db	1945	TGTAGAACCAAGGCGGATGTTTACTCTCTCCCGGATGGCAGTACCAACCATTAACCTTGAA	2004
Qy	1981	GTTAGCCCCCAAGAGAACCAACCGAGACCGCAGCTGACTGACAGATACAGGCCAAGTGTG	2040
Db	2005	GTTAGCCCCCAAGAGAACCAACCGAGACCGCAGCTGACTGACAGATACAGGCCAAGTGTG	2064

2041 GATTCTCAGGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2100
2065 GATTCTCAGGAAACCTTTAAATTTATCCGTGGGGCTCTTCACAAATGCGCGCAATT 2124
2101 TCTTTGAGTAAAGTTTGAATCCCAAGTATCACTGTCCATTTTGTTACCTTCAAGACATTT 2160
2125 TCTTTGAGTAAAGTTTGAATCCCAAGTATCACTGTCCATTTTGTTACCTTCAAGACATTT 2184
2161 TATCAGAAAGTTTGAATGATGACACAGAGACTGGAGCATAAATACAATCTCGAGTTTCAAT 2220
2185 TATCAGAAAGTTTGAATGATGACACAGAGACTGGAGCATAAATACAATCTCGAGTTTCAAT 2244
2221 AAAAAGTCTGAAACCAAGTCTTGTAGAGTCAAGTACCGGATGCGCGCAGCGTTG 2280
2245 AAAAAGTCTGAAACCAAGTCTTGTAGAGTCAAGTACCGGATGCGCGCAGCGTTG 2304
2281 CTGGGAAAGAGTGTCCCTCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTC 2340
2305 CTGGGAAAGAGTGTCCCTCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTC 2364
2341 CCGGCGAGTCAAAATCCCTGCTGCTGCGAAGGGGAAAGAGAGCCCTCTCGGCGCAGGC 2400
2365 CCGGCGAGTCAAAATCCCTGCTGCTGCGAAGGGGAAAGAGAGCCCTCTCGGCGCAGGC 2424
2401 AAGGCCCTCTGACTTTCAGGATAGACTCTAGCACTTTAGCCCAAGTAACTGAAGTCC 2460
2425 AAGGCCCTCTGACTTTCAGGATAGACTCTAGCACTTTAGCCCAAGTAACTGAAGTCC 2484
2461 CACAGCACACAGCAGAATGTGGGGTCCAAAGGGGCGCCACAGGCAACAGCAATCTGAG 2520
2485 CACAGCACACAGCAGAATGTGGGGTCCAAAGGGGCGCCACAGGCAACAGCAATCTGAG 2544
2521 ATGTTTCTTAAACCAAGTGTTCCTCTGCAACGGATAGACAAAGAGCCCGAGCAAAA 2580
2545 ATGTTTCTTAAACCAAGTGTTCCTCTGCAACGGATAGACAAAGAGCCCGAGCAAAA 2604
2581 TTGAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGGT 2640
2605 TTGAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCTCGGAGCAGTAAATCAATGGT 2664
2641 TCCATCGAGTACCCCGCAGCAAGCAGCAGCCGTCGAGCACTTCCGGGAGAGACTATTTTC 2700
2665 TCCATCGAGTACCCCGCAGCAAGCAGCAGCCGTCGAGCACTTCCGGGAGAGACTATTTTC 2724
2701 TGTAAATCGGAGTCCAGCAATATCTGACAGCAATTTGTTGAGCCCTTCCAAAAGACTG 2760
2725 TGTAAATCGGAGTCCAGCAATATCTGACAGCAATTTGTTGAGCCCTTCCAAAAGACTG 2784
2761 AAGTCCAGCGTGTGCTGCTTGAAGTGAACGAGCCCGGGGCAATTAACAGAGGCTAT 2820
2785 AAGTCCAGCGTGTGCTGCTTGAAGTGAACGAGCCCGGGGCAATTAACAGAGGCTAT 2844
2821 GACCTTCCCAAGTACCATATGTCAGAGGCAATCATCATCTGTTACCGCAGACTGTGTG 2880
2845 GACCTTCCCAAGTACCATATGTCAGAGGCAATCATCATCTGTTACCGCAGACTGTGTG 2904
2881 TATCGTCCGAGGCGTCCCTCCCAACCAAGTTTCTGAGCTCCAGCGAGTCTGATTCT 2940
2905 TATCGTCCGAGGCGTCCCTCCCAACCAAGTTTCTGAGCTCCAGCGAGTCTGATTCT 2964
2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTGTG 3000
2965 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGCTCCGGGCCACTTTACATTTGTGTG 3024
3001 CCTGCTGTGTCAGAGTCCAGCTCAGCTTGAAGGCTTGTGAGTGTGATGTCAGTCTTA 3060
3025 CCTGCTGTGTCAGAGTCCAGCTCAGCTTGAAGGCTTGTGAGTGTGATGTCAGTCTTA 3084
3061 CTCCCCCAAAATTTAAATTTTACTTCTATCTCTTGAAGAGCAATGGTGAAGACTACTGAA 3120
3085 CTCCCCCAAAATTTAAATTTTACTTCTATCTCTTGAAGAGCAATGGTGAAGACTACTGAA 3144

QY 3121 ATAACTCTGATTTCTGTACATATAACATATGAGGAATCTGCAAGCAACACTACAGTT 3180
Db 3145 ATAACTCTGATTTCTGTACATATAACATATGAGGAATCTGCAAGCAACACTACAGTT 3204
QY 3181 GTGTAA 3186
Db 3205 GTGTAA 3210
RESULT 2
AAV09024
ID AAV09024 standard; DNA; 3183 BP.
XX AAV09024;
AC AAV09024;
XX 21-JUL-1998 (first entry)
XX Homo sapiens 20q13 amplicon ZABC-1 cDNA sequence.
DE 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
KW treatment; age-related macular degeneration; retinitis pigmentation;
KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 1. 3183
FT /*tag= a
FT /product= "ZABC1 protein"
XX WO9802539-A1.
XX 22-JAN-1998.
XX 15-JUL-1997; 97WO-US012343.
XX 15-JUL-1996; 96US-00680395.
PR 16-OCT-1996; 96US-00731499.
PR 17-JAN-1997; 97US-00785532.
XX (REGC) UNIV CALIFORNIA.
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
WPI; 1998-110587/10.
DR P-PSDB; AAW23975.
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
PT abnormalities, particularly tumours, and for developing products for
PT treating diseases.
XX Claim 1; Page 67-69; 91pp; English.
XX The sequence is that of the cDNA sequence encoding ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
XX pigmentation
SQ Sequence 3183 BP; 925 A; 796 C; 758 G; 698 T; 0 U; 6 Other;

Query Match	99.1%;	Score 3156.4;	DB 2;	Length 3183;
Best Local Similarity	99.7%;	Prod. No. 0;		
Matches 3176;	Conservative	0;	Mismatches 7;	Indels 3; Gaps 1;
QY	1	ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTATCATCGATGGG	60	
DB	1	ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTATCATCGATGGG	60	
QY	61	CCAGAAGTGATGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG	120	
DB	61	CCAGAAGTGATGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG	117	
QY	121	AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATATCATCCAAATCGAG	180	
DB	118	AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATATCATCCAAATCGAG	177	
QY	181	GGGTATATGCCCTTGGATTGATGTTCTGAGCCAGACCTTCACATTCAGAAGACCTT	240	
DB	178	GGGTATATGCCCTTGGATTGATGTTCTGAGCCAGACCTTCACATTCAGAAGACCTT	237	
QY	241	AATAAACATGCTTAATGCAACACCGGCCTACCTCTGTGAACACAGACATCTTCGGGTT	300	
DB	238	AATAAACATGCTTAATGCAACACCGGCCTACCTCTGTGAACACAGACATCTTCGGGTT	297	
QY	301	GAAGCAGATATCTCAGTCCGCTGTATAAAGTCAAGTGGGAAACAGAACTCCCAAGAA	360	
DB	298	GAAGCAGATATCTCAGTCCGCTGTATAAAGTCAAGTGGGAAACAGAACTCCCAAGAA	357	
QY	361	AAGAAATTGCAAGGAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCCT	420	
DB	358	AAGAAATTGCAAGGAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCCT	417	
QY	421	TTTGATGTTGAGATCCACATCAGAAACACACAAAGATTCTTTCACTTACGGGTGTAAATG	480	
DB	418	TTTGATGTTGAGATCCACATCAGAAACACACAAAGATTCTTTCACTTACGGGTGTAAATG	477	
QY	481	TGCGGAAGAAGMTTSTRSSAGCCTTGGTTCTTTAAAAATCATATCGCGACACATAATGGC	540	
DB	478	TGCGGAAGAAGMTTSTRSSAGCCTTGGTTCTTTAAAAATCATATCGCGACACATAATGGC	537	
QY	541	AAATCGGGGGCCAGAAACAACTGACAGCAAGCCTTGGAGAGTAGTCCAGCAACGATCAAC	600	
DB	538	AAATCGGGGGCCAGAAACAACTGACAGCAAGCCTTGGAGAGTAGTCCAGCAACGATCAAC	597	
QY	601	GAGGTCGTCCAGGTGCACGCGCGAGAGCATCTCTCTCTTACAAAAATCTGCATGGTT	660	
DB	598	GAGGTCGTCCAGGTGCACGCGCGAGAGCATCTCTCTCTTACAAAAATCTGCATGGTT	657	
QY	661	TGTGGCTTCTATTTCAAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA	720	
DB	658	TGTGGCTTCTATTTCAAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA	717	
QY	721	AAACTGCTTTTCGTATCAGACAGCGCGACAGACTCTCACAAGGAGGATCCGCTCC	780	
DB	718	AAACTGCTTTTCGTATCAGACAGCGCGACAGACTCTCACAAGGAGGATCCGCTCC	777	
QY	781	TCGAGGGAGGACTTCTCTCAGTTGTTTCAACTTGAGACCAAAATCTCACCTTGAACCGGG	840	
DB	778	TCGAGGGAGGACTTCTCTCAGTTGTTTCAACTTGAGACCAAAATCTCACCTTGAACCGGG	837	
QY	841	AAGAAGCCTGTCCAGATGCATCCCTCAGCTCGATCCGTTTCCACCTTCCAGGCTTGGCAG	900	
DB	838	AAGAAGCCTGTCCAGATGCATCCCTCAGCTCGATCCGTTTCCACCTTCCAGGCTTGGCAG	897	
QY	901	CTGGCTACCAAGGAAAGTTGCCATTTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG	960	
DB	898	CTGGCTACCAAGGAAAGTTGCCATTTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG	957	
QY	961	AGCACCGACAACGACGATTTCCAGTTTCCGAGAAGGAGCTTTCGAGAAAAAATAAAGGCGAGT	1020	
DB	958	AGCACCGACAACGACGATTTCCAGTTTCCGAGAAGGAGCTTTCGAGAAAAAATAAAGGCGAGT	1017	
QY	1021	TGTGAGGCGCTCTCGCAAGAGAAAGAGAAAGTGCAAACTCCCCACGGCGAAGCGCCCTCC	1080	

Db 2098 TCTTTGAGTAAAGTTTGATTCCTCAAGTATCACCTGTCTCATTTTGTACCTTCAAGACATTT 2157
Qy 2161 TATCAGAAGTTTAAATGATGACACAGAGACTGGAGCATTAATACAATCTCGACGTTTCAAT 2220
Db 2158 TATCAGAAGTTTAAATGATGACACAGAGACTGGAGCATTAATACAATCTCGACGTTTCAAT 2217
Qy 2221 AAAAAGTGTGAAACCAAGTCTTGTGTAGAGTCCAGCTACCGGATGCCCGCAGCGTTG 2280
Db 2218 AAAAAGTGTGAAACCAAGTCTTGTGTAGAGTCCAGCTACCGGATGCCCGCAGCGTTG 2277
Qy 2281 CTGGAAAGAGTGTGCTCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Db 2278 CTGGAAAGAGTGTGCTCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2337
Qy 2341 CCGGCGAGTCCAAATCCCTGCCATCTCGGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC 2400
Db 2338 CCGGCGAGTCCAAATCCCTGCCATCTCGGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC 2397
Qy 2401 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTAGCCCAAGTAACCTGAAGTCC 2460
Db 2398 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTAGCCCAAGTAACCTGAAGTCC 2457
Qy 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2520
Db 2458 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2517
Qy 2521 ATGTTTCTAAACCAAGTGTTCCTCTGACCGGATAGCAAAAGACCCGAGACAAA 2580
Db 2518 ATGTTTCTAAACCAAGTGTTCCTCTGACCGGATAGCAAAAGACCCGAGACAAA 2577
Qy 2581 TTGAACCTCTCCAGTACTCTCTCTAGCCACCTCGGCGAGCAGTAACATCAATGTT 2640
Db 2578 TTGAACCTCTCTCCAGTACTCTCTCTAGCCACCTCGGCGAGCAGTAACATCAATGTT 2637
Qy 2641 TCCATCGACTACCCCGCCCAAGAACACAGCCGCGGCACTCTCGGGAAGAGACTATTTC 2700
Db 2638 TCCATCGACTACCCCGCCCAAGAACACAGCCGCGGCACTCTCGGGAAGAGACTATTTC 2697
Qy 2701 TGTAACTCGAGTGCAGCAATCTCTGAGCAAGATTTGTGTAGCCCTTCCAAAAGACTG 2760
Db 2698 TGTAACTCGAGTGCAGCAATCTCTGAGCAAGATTTGTGTAGCCCTTCCAAAAGACTG 2757
Qy 2761 AAGTCCAGGTGTGCTGCTTGTGAGTTCAGGTCAGCCCGGCGCAATACAGAGAGGCTAT 2820
Db 2758 AAGTCCAGGTGTGCTGCTTGTGAGTTCAGGTCAGCCCGGCGCAATACAGAGAGGCTAT 2817
Qy 2821 GACCTTCCCAAGTACCATATGTCAGAGCATCACATCACTGTTTACCAGAGACTGTGTG 2880
Db 2818 GACCTTCCCAAGTACCATATGTCAGAGCATCACATCACTGTTTACCAGAGACTGTGTG 2877
Qy 2881 TATCCGTCCGAGGCTGCTCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTTCT 2940
Db 2878 TATCCGTCCGAGGCTGCTCTCCCAACCAAGTTCTCTGAGCTCCAGCGAGTCTGATTTCT 2937
Qy 2941 CCAATGTGCTGACTGTTTCAAGGCCCTATGTTGCTCGGGCCACTTTACACTTGTGTG 3000
Db 2938 CCAATGTGCTGACTGTTTCAAGGCCCTATGTTGCTCGGGCCACTTTACACTTGTGTG 2997
Qy 3001 CTGCTGTGTAGTCCAGCATCCAGCTTCAGCTTGTGAGGTTTGTGAGTGTGAGTGTGCTTA 3060
Db 2998 CTGCTGTGTAGTCCAGCATCCAGCTTCAGCTTGTGAGGTTTGTGAGTGTGAGTGTGCTTA 3057
Qy 3061 CTCCCATGAAATTAATTTTATCTTCTCATCTTTTGAAGACCAATGGTGAAGACTACTGAA 3120
Db 3058 CTCCCATGAAATTAATTTTATCTTCTCATCTTTTGAAGACCAATGGTGAAGACTACTGAA 3117
Qy 3121 ATAAGCTGTGATGTACTGTACATTAACATATGAGGAATCTGCAAGGAACACTACAGTT 3180
Db 3118 ATAAGCTGTGATGTACTGTACATTAACATATGAGGAATCTGCAAGGAACACTACAGTT 3177
Qy 3181 GTGTAA 3186
Db 3178 GTGTAA 3183

RESULT 3

ID ACC50334 standard; cDNA; 5632 BP.

AC ACC50334;

DT 12-JUN-2003 (first entry)

Breast cancer associated cDNA sequence SEQ ID NO:505.

Human; breast cancer; cytostatic; gene therapy; gene; ss.

OS Homo sapiens.

PN WO2003004989-A2.

XX 16-JAN-2003.

XX 21-JUN-2002; 2002WO-US019669.

XX 21-JUN-2001; 2001US-0299887P.

PR 27-JUN-2001; 2001US-0301572P.

PR 18-JUL-2001; 2001US-0306501P.

PR 25-SEP-2001; 2001US-0325002P.

PR 05-MAR-2002; 2002US-0362585P.

PR 14-MAY-2002; 2002US-0380391P.

XX (MILL-) MILLENIUM PHARM INC.

XX Lillie J, Gannavarapu M, Glatt K, Hoerh S, Kamatkar S;

PI Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;

PI Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;

XX WPI; 2003-210381/20.

DR P-PSDB; ABR47632.

XX Breast cancer diagnosis or treatment by comparing the level of expression

of a marker in a patient sample with that in the control non-breast

cancer sample.

PS Claim 1; SEQ ID NO 505; 128pp; English.

XX The present invention describes a method for assessing whether a patient

is afflicted with breast cancer. The method comprises comparing the level

of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and

ABR47386 to ABR47632) in a patient sample and the normal level of

expression of the marker in a control non-breast cancer sample, where a

significant increase in the level of expression of the marker in the

patient sample and the normal level is an indication that the patient is

afflicted with breast cancer. The breast cancer associated sequences from

the present invention have cytostatic activities and can be used in gene

therapy. The method is useful for diagnosing and treating breast cancer.

N.B. the sequence data for this patent did not form part of the printed

specification, but was obtained in electronic format directly from WIPO

at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 95.4%; Score 3040.6; DB 8; Length 5632;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCAATCGAAGTGACAGGAACATGCCAACTCAATCCCTTAAATGTACATGATGG 60

Db 272 ATGCAATCGAAGTGACAGGAACATGCCAACTCAATCCCTTAAATGTACATGATGG 331

Qy 61 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG 120

Db 332 CCAGAAGTGATGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG 391

Qy 121 AAAGGGACCGCTGTGTGTTCCATTCGAGCTACACAGAAAAAATGTATCCAAATCGAG 180

Qy	2341	CCGGCGCAGTCCAAATCCCTGCCATCTGCAGAGGGGAAGCAGAGCCCTCTCTGGGCCAGGC	2400
Db	2612	CCGGCGCAGTCCAAATCCCTGCCATCTGCAGAGGGGAAGCAGAGCCCTCTCTGGGCCAGGC	2671
Qy	2401	AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACCTGGAAGTCC	2460
Db	2672	AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACCTGGAAGTCC	2731
Qy	2461	CACAGACCACAGCAGAAATGTGGGGTCCAAAGGGGCGCCACAGAGGCNAACAGCAATCTGAG	2520
Db	2732	CACAGACCACAGCAGAAATGTGGGGTCCAAAGGGGCGCCACAGAGGCNAACAGCAATCTGAG	2791
Qy	2521	ATGTTTCTTAAACACAGTGTTCCTCCCTGCACCGATAAGACAATAAGACCCGAGACAAA	2580
Db	2792	ATGTTTCTTAAACACAGTGTTCCTCCCTGCACCGATAAGACAATAAGACCCGAGACAAA	2851
Qy	2581	TTGAACCTCTTCCAGTAGTCTCTTTCAGCCCAACCTTCGGCGACGATTAACATCAATGGT	2640
Db	2852	TTGAACCTCTTCCAGTAGTCTCTTTCAGCCCAACCTTCGGCGACGATTAACATCAATGGT	2911
Qy	2641	TCCATCGACTACCCCGCAAGACGACAGCCCTGGGCACTCTCGGGAAGAGACTATTTTC	2700
Db	2912	TCCATCGACTACCCCGCAAGACGACAGCCCTGGGCACTCTCGGGAAGAGACTATTTTC	2971
Qy	2701	TGTAATCGGAGTGCAGCAATACTGCGACAGAAATTTGGTGAGCCCTTCCAAAAAGACTG	2760
Db	2972	TGTAATCGGAGTGCAGCAATACTGCGACAGAAATTTGGTGAGCCCTTCCAAAAAGACTG	3031
Qy	2761	AAGTCCAGCGTGTTCCTTGAGCTTGACAGCCCGGGGCCAATTACAGAGAGCGTAT	2820
Db	3032	AAGTCCAGCGTGTTCCTTGAGCTTGACAGCCCGGGGCCAATTACAGAGAGCGTAT	3091
Qy	2821	GACCTTCCCAAGTACCATAATGGTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG	2880
Db	3092	GACCTTCCCAAGTACCATAATGGTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG	3151
Qy	2881	TATCCGTCGAGCGCTGCTCCCAACCAAGGTTCTTGAGCTCCAGCAGGTCGATTCCT	2940
Db	3152	TATCCGTCGAGCGCTGCTCCCAACCAAGGTTCTTGAGCTCCAGCAGGTCGATTCCT	3211
Qy	2941	CAAAATGCTGCTGACTGTTCAAGAGCCCTATGTGGCTCCGGGCCACTTTTACACTGTGTG	3000
Db	3212	CAAAATGCTGCTGACTGTTCAAGAGCCCTATGTGGCTCCGGGCCACTTTTACACTGTGTG	3271
Qy	3001	CTTGCTGGTAGTCCAGCATCCAGCTCGACTTAG-----	3034
Db	3272	CTTGCTGGTAGTCCAGCATCCAGCTCGACTTAGAGGAAAAAGGCCTGTGTCATATCAA	3331
Qy	3035	-----	3034
Db	3332	CACTTATCTAACAGCATGGCAAAAGAGAAACTATGAGAAATTTTATTTGGGAATGCACAT	3391
Qy	3035	-----	3047
Db	3392	TATCGACAAAATGACAAAAAACTTGATTCACTAATTTAGGGGGAATAAGTCTTGTGTG	3451
Qy	3048	ATGTCAGTCTTACTCCCAATGAAATTAATTTTACATTCTTTGAGAGCGAATGGT	3107
Db	3452	ATGTCAGTCTTACTCCCAATGAAATTTAATTTTACATTCTTTGAGAGCGAATGGT	3511
Qy	3108	GAAGGCTACTGAAATAGCTGTGATTGTACTGTACATAAAAAATATGAGGAATCTGCAAG	3167
Db	3512	GAAGGCTACTGAAATAGCTGTGATTGTACTGTACATAAAAAATATGAGGAATCTGCAAG	3571
Qy	3168	GAACACTACAGTTGTGTAA	3186
Db	3572	GAACACTACAGTTGTGTAA	3590

RESULT 4

ADD14635

ID ADD14635 standard; cDNA; 5632 BP.

	ADD14635;	
XX	AC	XX
XX	DT	XX
XX	DE	XX
XX	PI	Human src biomarker polynucleotide SEQ ID NO:29.
XX	KW	predictor set; protein tyrosine kinase activity modulator;
XX	KW	protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
XX	KW	gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
OS	Homo sapiens.	
XX	PN	WO2003062395-A2.
XX	PD	31-JUL-2003.
XX	PP	17-JAN-2003; 2003WO-US001981.
XX	PR	18-JAN-2002; 2002US-0350061P.
XX	PA	(BRIM) BRISTOL-MYERS SQUIBB CO.
XX	PI	Huang F, Fairchild CR, Lee FY, Shaw P;
XX	DR	WPI; 2003-636735/60.
XX	DR	P-PSDB; ADD14040.
PT	New polynucleotides and polypeptides for predicting the activity of compounds that interact with protein tyrosine kinases and/or protein tyrosine kinase pathways.	
XX	Claim 2; SEQ ID NO 29; 139pp; English.	
CC	The present invention describes a predictor set comprising a plurality of polynucleotides or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity or members of the protein tyrosine kinase pathway. Also described: (1) predicting whether a compound is capable of modulating the activity of cells, comprising obtaining a sample of cells, determining whether the cells express a plurality of markers, and correlating the expression of the markers to the compound's ability to modulate the activity of the cells; (2) a plurality of cell lines for identifying polynucleotides and polypeptides whose expression levels correlate with compound sensitivity or resistance of cells associated with a disease state; and (3) identifying polynucleotides and polypeptides that predict compound sensitivity or resistance of cells associated with a disease state, comprising subjecting the plurality of cell lines to one or more compounds, analysing the expression pattern of a microarray of polynucleotides or polypeptides, and selecting polynucleotides or polypeptides that predict the sensitivity or resistance of cells associated with a disease state by using the expression pattern of the microarray. The polynucleotides and polypeptides have cytoskeletal activities, and can be used in gene therapy. The polynucleotides and polypeptides are useful in predicting the activity of compounds that interact with protein tyrosine kinases and/or protein tyrosine kinase pathways. These may be used in determining drug sensitivity in patients to allow the development of individualized genetic profiles which aid in treating diseases and disorders (e.g. cancer) based on patient response at a molecular level. The present sequence is used in the exemplification of the present invention.	
XX	Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;	
SQ	Query Match 95.4%; Score 3040.6; DB 10; Length 5632; Best Local Similarity 95.8%; Pred. No. 0; Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;	
Oy	1 ATGCAATCGAAAGTGCAGGAACAATGCCAATCCTTCCTTTTAATGTACATGGATGGG 60 	
Dd	272 ATGCAATCGAAAGTGCAGGAACAATGCCAATCCTTCCTTTTAATGTACATGGATGGG 331 	
Oy	61 CCAGAAGTGATTGGCAGCTCTCTTTGGCAGTCCGATGGAGATGGAGATGCTTTGTCAATG 120	

XX	Sequence	5632 BP; 1637 A; 1167 C; 1257 G; 1571 T; 0 U; 0 Other;	
SQ	Query Match	95.4%; Score 3040.6; DB 13; Length 5632;	
	Best Local Similarity	95.8%; Pred. No. 0;	
	Matches 3180; Conservative	6; Mismatches 0; Indels 133; Gaps 1;	
Qy	1	ATGCAATCGAAAGTGACAGGAAAATGCGCAACTCAATCCCTCTTAATGTACATGGATGGG	60
Db	272	ATGCAATCGAAAGTGACAGGAAAATGCGCAACTCAATCCCTCTTAATGTACATGGATGGG	331
Qy	61	CCAAAGTGATGGGAGCTCTCTTGGGAGTCCGATGGAGATGGAGATGCTTGTCAATG	120
Db	332	CCAAAGTGATGGGAGCTCTCTTGGGAGTCCGATGGAGATGGAGATGCTTGTCAATG	391
Qy	121	AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCATCCAAATCGAG	180
Db	392	AAAGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCATCCAAATCGAG	451
Qy	181	GGGTATATGCCCTTGGATGTGATGTTCTGACGCCAGACCTTCAACATTCAGAAAGACCTT	240
Db	452	GGGTATATGCCCTTGGATGTGATGTTCTGACGCCAGACCTTCAACATTCAGAAAGACCTT	511
Qy	241	AATAAATGCTTTAATGCAACACCGGCTTACCTCTGTGAACGACGAGTTCTTCCGGGTT	300
Db	512	AATAAATGCTTTAATGCAACACCGGCTTACCTCTGTGAACGACGAGTTCTTCCGGGTT	571
Qy	301	GAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGACCTCCCAAGGAA	360
Db	572	GAACGAGATATCTCAGTCCGCTTGATAAAGTCAAGTGCAGACAGACCTCCCAAGGAA	631
Qy	361	AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTGGCT	420
Db	632	AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTGGCT	691
Qy	421	TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAAATG	480
Db	692	TTTGATGTTGAGATCCACATGAGAACACACAAAGATTTCTTCACTTACGGGTGAAATG	751
Qy	481	TGCGGAAGAGMTTSRRSSAGCCCTGGTTTCTTAAATATCATGCGGACACATTAATGGC	540
Db	752	TGCGGAAGAGMTTCAAGAGCCCTGGTTTCTTAAATATCATGCGGACACATTAATGGC	811
Qy	541	AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC	600
Db	812	AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAAC	871
Qy	601	GAGTGTGTCAGGTGCAACGGGCGGAGAGCATCTCTCTCTTACAAATCTGCATGGTT	660
Db	872	GAGTGTGTCAGGTGCAACGGGCGGAGAGCATCTCTCTCTTACAAATCTGCATGGTT	931
Qy	661	TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGACCGCAAGGTGCACACAAA	720
Db	932	TGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGACCGCAAGGTGCACACAAA	991
Qy	721	AAACTGCTTTCCGTATCCAGACGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC	780
Db	992	AAACTGCTTTCCGTATCCAGACGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCC	1051
Qy	781	TCGAGGGAGGACTTCCCTCAGTGTGTTCAACTTGGAGACCAAAATCTCACCTTGAAACGGGG	840
Db	1052	TCGAGGGAGGACTTCCCTCAGTGTGTTCAACTTGGAGACCAAAATCTCACCTTGAAACGGGG	1111
Qy	841	AAGAGCCTGTCAAGTATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG	900
Db	1112	AAGAGCCTGTCAAGTATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG	1171
Qy	901	CTGGCTACCAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAAGGG	960
Db	1172	CTGGCTACCAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAAGGG	1231
Qy	961	AGCACCGACACGACGATTCGAGTTCCGAGTAAGGAGCTTGGAGAAAACAAATAAGGGCAGT	1020

Db	1232	AGCACCGACAAACGACGATTCGAGTTCCGAGAAAGAGCTTGGAGAAAACAAATAAGGGCAGT	1291
Qy	1021	TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACACTCCCACGGCGAAGCCCTCC	1080
Db	1292	TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACACTCCCACGGCGAAGCCCTCC	1351
Qy	1081	GTGACGCGGATCCAAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGC	1140
Db	1352	GTGACGCGGATCCAAAGTTACCCAGTAGCAAGGAGAACCCACTCACTGCTCCGAGTGC	1411
Qy	1141	GGCAAGCTTTACAGAACTTACCACCGCTGGTCTTGCACTCCAGGGTCCACAAGAGGAC	1200
Db	1412	GGCAAGCTTTACAGAACTTACCACCGCTGGTCTTGCACTCCAGGGTCCACAAGAGGAC	1471
Qy	1201	CGAGGGCCGCGCGGAGTCCGCCACCATGTCTGTGACCGGAGGAGCGGGGACGTGT	1260
Db	1472	CGAGGGCCGCGCGGAGTCCGCCACCATGTCTGTGACCGGAGGAGCGGGGACGTGT	1531
Qy	1261	TCTCTGACCTCGCCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAAGGTGGT	1320
Db	1532	TCTCTGACCTCGCCGCCCTCTGGATGAAATGGAGCCGTGGATCGAGGGGAAAGGTGGT	1591
Qy	1321	TCTGAACGCGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAT	1380
Db	1592	TCTGAACGCGATCTGAGGATGGCTTCCCGAAGGAATCCATCTGGATAAAAATGATGAT	1651
Qy	1381	GGAGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAAGTTTTTC	1440
Db	1652	GGAGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGGAAAAGTTTTTC	1711
Qy	1441	CGTTCAAATTTATACCTCAATATTCATCTCAGAACCGATACAGGTGAAAAACCATACAAA	1500
Db	1712	CGTTCAAATTTATACCTCAATATTCATCTCAGAACCGATACAGGTGAAAAACCATACAAA	1771
Qy	1501	TGTGAATTTTGTGAATATGCTGCAGCCGAGAGACATCTCTGAGGTATCACTTCGAGAGA	1560
Db	1772	TGTGAATTTTGTGAATATGCTGCAGCCGAGAGACATCTCTGAGGTATCACTTCGAGAGA	1831
Qy	1561	CATCAAGGAAAAACAAACCGATGTTGCTGCTCAAGTCAAGAACGATGGTAAAAATCAG	1620
Db	1832	CATCAAGGAAAAACAAACCGATGTTGCTGCTCAAGTCAAGAACGATGGTAAAAATCAG	1891
Qy	1621	GACACTGAAGTGCATTTAAACCGCTGACAGTGGCGCAACCAAAAATTTGAAAAGATTT	1680
Db	1892	GACACTGAAGTGCATTTAAACCGCTGACAGTGGCGCAACCAAAAATTTGAAAAGATTT	1951
Qy	1681	TTTGATGGTCCAAAGATGTTACAGGAGTCCACCTGCAGAGAGCTTAAAGAGATGCT	1740
Db	1952	TTTGATGGTCCAAAGATGTTACAGGAGTCCACCTGCAGAGAGCTTAAAGAGATGCT	2011
Qy	1741	TCTGTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACACGACACAAAGATACCTCAG	1800
Db	2012	TCTGTTTTTCAGAAATGTTCTGGGCGAGCTGCTCTCTCACACGACACAAAGATACCTCAG	2071
Qy	1801	GATTTCCATAAAAAATGCAGCTGATGACAGTCTGATTAAGTGAATATAAAACCTTACCCCT	1860
Db	2072	GATTTCCATAAAAAATGCAGCTGATGACAGTCTGATTAAGTGAATATAAAACCTTACCCCT	2131
Qy	1861	GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATAACCTCATC	1920
Db	2132	GCTTACCTGGACCTGTTAAAAAAGAGATCAGCAGTTGAAAACCTCAGGCAAAATAACCTCATC	2191
Qy	1921	TGTAGAACCAAGGGGATGTTTACTCTCTCCGATGGCAGTACCCACCATTAACCTTGAA	1980
Db	2192	TGTAGAACCAAGGGGATGTTTACTCTCTCCGATGGCAGTACCCACCATTAACCTTGAA	2251
Qy	1981	GTTAGCCCAAGAGAGACAAACCGAGACCGGAGCTGATCGAGATACAGGCAAGTGTG	2040
Db	2252	GTTAGCCCAAGAGAGACAAACCGAGACCGGAGCTGATCGAGATACAGGCAAGTGTG	2311
Qy	2041	GATTTGTCAGAAAAACCTTTAAATTTATCCGTGGGGCTCTTCAAAATTTCCCGCAATT	2100
Db	2312	GATTTGTCAGAAAAACCTTTAAATTTATCCCGTGGGGCTCTTCAAAATTTCCCGCAATT	2371

QY 2101 TCTTTGAGTAAAGTTTGAATCCAGATATCACCTGTCATTTTGTACCTTCAAGACATTT 2160
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2372 TCTTTGAGTAAAGTTTGAATCCAGATATCACCTGTCATTTTGTACCTTCAAGACATTT 2431
QY 2161 TATCAGAGTTTAAATGATGACACAGAGACTGGAGCATATAAATCAATCTCGACGTTTCAAT 2220
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2432 TATCAGAGTTTAAATGATGACACAGAGACTGGAGCATATAAATCAATCTCGACGTTTCAAT 2491
QY 2221 AAAAATGTGCAAAAGTCTTGTCTGATAGAGTCAAGTACCGGATGCCCGCAGCGTTG 2280
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2492 AAAAATGTGCAAAAGTCTTGTCTGATAGAGTCAAGTACCGGATGCCCGCAGCGTTG 2551
QY 2281 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2340
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2552 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTC 2611
QY 2341 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTGGGCGCAGGC 2400
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2612 CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTGGGCGCAGGC 2671
QY 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCC 2460
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACCTGAAGTCC 2731
QY 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGCAACAGCAATCTGAG 2520
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2732 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGCAACAGCAATCTGAG 2791
QY 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAGAACCCGAGACAAA 2580
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2792 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAGAACCCGAGACAAA 2851
QY 2581 TTGAACCTCTTCCAGTACTCTTCTGAGCCCACTCTGGCAGCAGTAACATCAATGCT 2640
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2852 TTGAACCTCTTCCAGTACTCTTCTGAGCCCACTCTGGCAGCAGTAACATCAATGCT 2911
QY 2641 TCCATCGATACCCCGCCAGAACACACAGCCGTCGGCACCTCCGGGAAGAGACTATTTC 2700
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2912 TCCATCGATACCCCGCCAGAACACACAGCCGTCGGCACCTCCGGGAAGAGACTATTTC 2971
QY 2701 TGTATCGAGTCCAGCAATACTGACGAGAAATTTGTTGAGCCCTCTCCAAAAGAGCTG 2760
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2972 TGTATCGAGTCCAGCAATACTGACGAGAAATTTGTTGAGCCCTCTCCAAAAGAGCTG 3031
QY 2761 AAGTCCAGCGTGTGTCCTTGTGACGTTGACAGCCCGGGCCCAATTACAGAGAGGCTAT 2820
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3032 AAGTCCAGCGTGTGTCCTTGTGACGTTGACAGCCCGGGCCCAATTACAGAGAGGCTAT 3091
QY 2821 GACCTTCCAGTACATATGTCAGAGCATCATCATCTGTTTACCGAGGACTGTGTG 2880
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3092 GACCTTCCAGTACATATGTCAGAGCATCATCATCTGTTTACCGAGGACTGTGTG 3151
QY 2881 TATCGTCCAGCGCTGCTCCCAAAACCAAGGTTTCTGAGCTCCAGCGAGTCCGATTCT 2940
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3152 TATCGTCCAGCGCTGCTCCCAAAACCAAGGTTTCTGAGCTCCAGCGAGTCCGATTCT 3211
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3000
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3212 CCAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3271
QY 3001 CCGTGTGTGAGTCCAGCATTCAGCTCGACGTTAG----- 3034
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3272 CCGTGTGTGAGTCCAGCATTCAGCTCGACGTTAGAGGAAAGAGGCTGTGTGTCATCAATCA 3331
QY 3035 ----- 3034
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3332 CACTTATCTAACAGCATGTCACAAAGAGAAACTATGAGAAATTTATTGGGAATGCAAT 3391
QY 3035 ----- 3034
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3392 TATGACCAATGACAAAAAACTGATTCATTAATTAGGGGAAAAAGAGTCTTGGTGG 3451

QY 3048 ATGTCAGTGTCTACTCCCATGAAATTAATTTTACTTTCATCCTTTGAGAGCGAATGCT 3107
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3452 ATGTCAGTGTCTACTCCCATGAAATTAATTTTACTTTCATCCTTTGAGAGCGAATGCT 3511
QY 3108 GAAAGCTACTGAAATAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3167
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3512 GAAAGCTACTGAAATAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
QY 3168 GAAACTACAGTTGTGTAA 3186
DB |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
3572 GAACACTACAGTTGTGTAA 3590

RESULT 6
ADR65858
ID ADR65858 standard; DNA; 5632 BP.
XX
AC ADR65858;
XX
DT 02-DEC-2004 (first entry)
XX
DE Human prostatic carcinoma derived DNA SEQ ID 54 #1.
XX
KW human; cytostatic; diagnosis; prostatic cancer;
KW differential expression analysis; ds.
XX
OS Homo sapiens.
XX
PN WO2004076614-A2.
XX
PD 10-SEP-2004.
XX
PF 22-FEB-2004; 2004WO-DE000433.
XX
PR 17-FEB-2003; 2003DE-01009985.
PR 14-MAY-2003; 2003DE-01022134.
XX
XX (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (ROSE/) ROSENTHAL A.
PA (HERM/) HERMANN K.
PA (PILA/) PILARSKY C.
XX
Hinzmann B, Dahl E, Rosenthal A, Hermann K, Pilarczyk C, Specht T;
Schmitt A, Beckmann G, Bruemendorf T, Kinnemann H, Roepcke S;
Xinzhong L, Staub E;
WPI; 2004-653386/63.
XX
New nucleic acids, and encoded proteins, from prostatic cancer tissue,
useful for diagnosis, treatment and in screening for specific binding
agents.
XX
Claim 1; Page 253-254; 1607pp; German.
XX
This invention describes novel cytostatic polynucleotide and polypeptide
sequences which can be used in a method for diagnosing prostatic cancer
or the risk of developing prostatic cancer. Diagnosis is based on
determining over transcription or over expression of the sequences in
prostatic tissue. Screening for inhibitors of the sequences or detection
of substances involves a binding assay, any compounds that bind are
selected, optionally after deconvolution of mixtures. Detection of a
predetermined minimum level of the reporter indicates the presence of
tumour cells. Inhibitors can be chosen from antisense oligonucleotides,
short-interfering RNA or ribozymes; an organic molecule of molecular
weight below 5000, preferably 300, that binds to the polypeptide; an
apamer against the polypeptide; a (monoclonal) antibody (Ab) against the
polypeptide, preferably humanised or human; an anti-idiotypic, non-human
(monoclonal) antibody directed against Ab or any of the above derivatised
with a reporter group, cell toxin, immunostimulatory molecules and/or
radioisotope. The polynucleotides are identified in human prostatic
cancer by differential expression analysis, using DNA microarrays,
between normal and tumorous tissues, with (over)expression being detected

CC by quantitative PCR. Analysis of prostatic cancer samples showed that
CC CD24 was upregulated in many of them. Sections of tissue, isolated from
CC prostatic cancer patients, or subjects at risk, were incubated
CC sequentially with anti-human CD4 murine monoclonal antibodies;
CC biotinylated second antibody; streptavidin-conjugated horseradish
CC peroxidase and then diaminobenzidine as colour former (brown). The
CC samples were counterstained with hemalum (blue). Malignant cells stained
CC strongly but non-malignant cells only weakly. In 15 of 63 samples of
CC adenocarcinoma, membrane and cytoplasmic staining was very strong, and
CC lymph node metastases were also stained. ADR65805-ADR66954 represent the
CC polynucleotide and polypeptide sequences used in the method of the
CC invention.

XX
SQ Sequence 5632 BP; 1637 A; 1166 C; 1258 G; 1571 T; 0 U; 0 Other;

Query Match 95.4%; Score 3040.6; DB 13; Length 5632;
Best Local Similarity 95.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 60
Db 272 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG 331

Qy 61 CCAGAAGTGATGGCAGCTCTCTGGCAGTCGGATCGATGGAGATGGCCTTGTCAATG 120
Db 332 CCAGAAGTGATGGCAGCTCTCTGGCAGTCGGATCGATGGAGATGGCCTTGTCAATG 391

Qy 121 AAAGGGACCGCTGTGTTCATTCGGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 180
Db 392 AAAGGGACCGCTGTGTTCATTCGGAGCTACACAAGAAAAAATGTCCATCCAAATCGAG 451

Qy 181 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGACCTT 240
Db 452 GGGTATATGCCCTTGGATTTGCATGTTCTGCAGCCAGACCTTTCACACATTCAGAAAGACCTT 511

Qy 241 AATAAAATGCTTAAATGCAACACCGGCTACCTCTGTGAAACCGAGCTTCTTCGGGTT 300
Db 512 AATAAACATGCTTAAATGCAACACCGGCTACCTCTGTGAAACCGAGCTTCTTCGGGTT 571

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGACCTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGTGCAGAACAGACCTCCCAAGGAA 631

Qy 361 AAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCT 420
Db 632 AAGAAATGCAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCT 691

Qy 421 TTTGATGTTGAGATCCATGAGAAACACAAAAGATTCTTTTCACTTACGGGTGTAAACATG 480
Db 692 TTTGATGTTGAGATCCATGAGAAACACAAAAGATTCTTTTCACTTACGGGTGTAAACATG 751

Qy 481 TGCGAAGCAAGMTTSRRSSAGCCTTGGTTTCTTAAATAATCATGCGGACACATAATGCG 540
Db 752 TGCGAAGCAAGATTCAAGGAGCCTTGGTTTCTTAAATAATCATGCGGACACATAATGCG 811

Qy 541 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTTGGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGAACTGCAGCAAGGCTTTGGAGTAGTCCAGCAACGATCAAC 871

Qy 601 GAGTGTGTCAGGTGCAACGGGCCGAGAGCATCTCTCTCTTCAAAAATCTGCATGGTT 660
Db 872 GAGTGTGTCAGGTGCAACGGGCCGAGAGCATCTCTCTCTTCAAAAATCTGCATGGTT 931

Qy 661 TGTGCTTCTTATTTTCCAAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACCAA 720
Db 932 TGTGCTTCTTATTTTCCAAAATAAGAAAGTCTAATTGAGCACCGCAAGTGCACACCAA 991

Qy 721 AAAATGCTTTTCGGTACCAGACGCGGACAGACAGACTCTCCACAGGAGGAATGCCGTC 780
Db 992 AAAATGCTTTTCGGTACCAGACGCGGACAGACAGACTCTCCACAGGAGGAATGCCGTC 1051

Qy 781 TCGAGGAGGACTTCTCGAGTTGTTCAAATTGAGACCAAAATCTCACCTCGAAACGGGG 840
Db 940 TCGAGGAGGACTTCTCGAGTTGTTCAAATTGAGACCAAAATCTCACCTCGAAACGGGG 940

Db 1052 TCGAGGAGGACTTCTCGCAGTTGTTCAAATTGAGACCAAAATCTCACCTCGAAACGGGG 1111

Qy 841 AAGAAGCTGTGATGATCATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900

Db 1112 AAGAAGCTGTGATGATCATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 1171

Qy 901 CTGGCTACCAAGAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGCAAGAAGGG 960

Db 1172 CTGGCTACCAAGAAAGTTGCCATTTGCCAAGAAGTGAAGGAATCGGGCAAGAAGGG 1231

Qy 961 AGCACCCACACGACGATTCGAGTTCCGAGAAGGAGCTTCGAGAAACAAAATAAGCGCAGT 1020

Db 1232 AGCACCCACACGACGATTCGAGTTCCGAGAAGGAGCTTCGAGAAACAAAATAAGCGCAGT 1291

Qy 1021 TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACATCCCAACGGCGAAGCCCTCC 1080

Db 1292 TGTGACGGCTCTCGCAAGAGAAAGAGAGTGCAAAACATCCCAACGGCGAAGCCCTCC 1351

Qy 1081 GTGACGCGGATCCAAAGTTTACCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC 1140

Db 1352 GTGACGCGGATCCAAAGTTTACCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC 1411

Qy 1141 GGCAAAGCTTTCAGAACTTACCACAGCTGTCTTGCACTCCAGGGTCCACAAGAGGAC 1200

Db 1412 GGCAAAGCTTTCAGAACTTACCACAGCTGTCTTGCACTCCAGGGTCCACAAGAGGAC 1471

Qy 1201 CGAGAGGCGGCGGAGTGCAGCCACCATGTCTGTGACGGGAGGAGCCGGGACCTGT 1260

Db 1472 CGAGAGGCGGCGGAGTGCAGCCACCATGTCTGTGACGGGAGGAGCCGGGACCTGT 1531

Qy 1261 TCTCCTGACCTCGCGCCCTCTGGATGAAATGGAGCGCTGGATTCGAGGGGAAGTGGT 1320

Db 1532 TCTCCTGACCTCGCGCCCTCTGGATGAAATGGAGCGCTGGATTCGAGGGGAAGTGGT 1591

Qy 1321 TCTCAAGACGGATCTGAGGATGGCTTCCGAAAGGAATCCATCTGGATAAAAATGATGAT 1380

Db 1592 TCTCAAGACGGATCTGAGGATGGCTTCCGAAAGGAATCCATCTGGATAAAAATGATGAT 1651

Qy 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAAGTTTTTC 1440

Db 1652 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAAGTTTTTC 1711

Qy 1441 CGTTCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1500

Db 1712 CGTTCAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1771

Qy 1501 TGTCAATTTTGTGAATATGCTGCAGCCAGAAAGACATCTCTGAGGTATCACTTGGAGAGA 1560

Db 1772 TGTCAATTTTGTGAATATGCTGCAGCCAGAAAGACATCTCTGAGGTATCACTTGGAGAGA 1831

Qy 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAAACGATGGTAAAAATCAG 1620

Db 1832 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAAGTCAAGAAACGATGGTAAAAATCAG 1891

Qy 1621 GACATGAAGATGACATTAATTAACCGCTGACAGTGCAGAAACCAAAAATTTGAAAAGATTT 1680

Db 1892 GACATGAAGATGACATTAATTAACCGCTGACAGTGCAGAAACCAAAAATTTGAAAAGATTT 1951

Qy 1681 TTTGATGTTGCAAGAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAGAGAGATGCT 1740

Db 1952 TTTGATGTTGCAAGAGATGTTTACAGGAGTCCACTGCAAGAGAGCTTAAGAGAGATGCT 2011

Qy 1741 TCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCTCACGAGCACACAAAGATCTCAG 1800

Db 2012 TCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCTCACGAGCACACAAAGATCTCAG 2071

Qy 1801 GATTTCATAAAAATGACGCTGATGACAGTCTGATTAAGTGAATAAAAAATTTTCCCTTCCCT 1860

Db 2072 GATTTCATAAAAATGACGCTGATGACAGTCTGATTAAGTGAATAAAAAATTTTCCCTTCCCT 2131

Qy 1861 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAATCTCAGGCAAAATAAATCTCATC 1920

Db 2132 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAATCTCAGGCAAAATAAATCTCATC 2191

CC with a bisulfite reagent, to convert non-methylated C to uracil or some
CC other base having base-pairing properties different from those of C.
CC Fragments of the treated DNA are then amplified, using primers and a
CC polymerase and the methylation status of the genomic CpG dinucleotides is
CC determined by analysis of the amplicons, particularly by hybridisation to
CC the oligomer, optionally with extension of the hybridised oligomer by at
CC least one base, or detection is by sequencing. The amplification may use
CC methylation-specific primers. Alternatively, genomic DNA is extracted
CC from a sample and digested with methylation-specific restriction enzymes,
CC then the digestion fragments detected, optionally after amplification. In
CC either method, more than 10 fragments of 100-200 bp are amplified in a
CC single reaction vessel, using a heat-resistant DNA polymerase in PCR. The
CC amplicons carry detectable markers, e.g. fluorophores, radioisotopes
CC and/or releasable fragments of known mass that can be detected by mass
CC spectrometry. The method is used for characterisation, classification,
CC differentiation, staging, phase-estimation, diagnosis and/or therapy of
CC proliferative diseases of breast cells. The method provides very specific
CC classification of proliferative diseases, allowing better treatment. It
CC can both characterise methylation status and detect single-nucleotide
CC polymorphisms. This sequence represents human gene used to illustrate the
CC method of the invention. NOTE: This sequence does not appear in the
CC printed specification but has been retrieved from Genbank.
xx

Sequence 5653 BP; 1656 A; 1167 C; 1258 G; 1572 T; 0 U; 0 Other;

Query Match 95.4%; Score 3040.6; DB 12; Length 5653;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCAATCGAAGTGACAGAAACATGCCAACTCAATCCCTCTTAATGTAATGATGGG 60
Db 272 ATGCAATCGAAGTGACAGAAACATGCCAACTCAATCCCTCTTAATGTAATGATGGG 331
Qy 61 CCAGAGTGATGGCAGCTCTCTGGCAGTCGGATCGATGGAGATGGAGTGCCTTGTCAATG 120
Db 332 CCAGAGTGATGGCAGCTCTCTGGCAGTCGGATCGATGGAGATGGAGTGCCTTGTCAATG 391
Qy 121 AAAGGGACCGCTGTGTTCCATTCGAGCTACACAAAGAAAAAATGTCCATCCAAATCGAG 180
Db 392 AAAGGGACCGCTGTGTTCCATTCGAGCTACACAAAGAAAAAATGTCCATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTTGATGTTCTGACGCCAGACCTTTCACACATTCAGAGACCTT 240
Db 452 GGGTATATGCCCTTGGATTTGATGTTCTGACGCCAGACCTTTCACACATTCAGAGACCTT 511
Qy 241 AATAAACATGCTTAAATGCAACACCGGCTACCCCTCTGTGAACGAGAGTCTTCGGGTT 300
Db 512 AATAAACATGCTTAAATGCAACACCGGCTACCCCTCTGTGAACGAGAGTCTTCGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAA 631
Qy 361 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 420
Db 632 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTCGCT 691
Qy 421 TTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAACATG 480
Db 692 TTGATGTTGAGATCCATGAGAACACACAAAGATTTCTTTCACTTACGGGTGTAACATG 751
Qy 481 TCGGGAAGAGMTTSRRSSAGCCTTGGTTTCTTTAAAAATTCATCGCGACACATAATGGC 540
Db 752 TCGGGAAGAGMTTCAAGAGCCTTGGTTTCTTTAAAAATTCATCGCGACACATAATGGC 811
Qy 541 AAATCGGGGGCCAGAGCAAACTGACGAAGCTTTGGAGGTGAGTGCACAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAGCAAACTGACGAAGCTTTGGAGGTGAGTGCACAGCAACGATCAAC 871
Qy 601 GAGGTGCTCCAGGTGACCGGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGGTT 660
Db 872 GAGGTGCTCCAGGTGACCGGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGGTT 931

Qy 661 TGTGGCTTCCCTATTTCCTCAAAATAAAGTCTTAATGAGCACCGCAAGGTGCAACCAAA 720
Db 932 TGTGGCTTCCCTATTTCCTCAAAATAAAGTCTTAATGAGCACCGCAAGGTGCAACCAAA 991
Qy 721 AAAATCTGTTTCGGTATCCAGCAGCGCGAGACAGACTCTCCAAAGAGGAAATGCCGTCC 780
Db 992 AAAATCTGTTTCGGTATCCAGCAGCGCGAGACAGACTCTCCAAAGAGGAAATGCCGTCC 1051
Qy 781 TCGAGGAGGAGACTTCCCTGCGAGTTCTCAACTTGTAGACCAAAATCTCAACCTGAAACGGG 840
Db 1052 TCGAGGAGGAGACTTCCCTGCGAGTTCTCAACTTGTAGACCAAAATCTCAACCTGAAACGGG 1111
Qy 841 AAGAAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 900
Db 1112 AAGAAGCTGTGCAGATGCATCCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCAG 1171
Qy 901 CTGCTTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAGGG 960
Db 1172 CTGCTTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAGGG 1231
Qy 961 AGCACCGACACGACGATTTCGAGTTCCGAGAGGAGCTTTGGAGAAACAAATAAGGGCAGT 1020
Db 1232 AGCACCGACACGACGATTTCGAGTTCCGAGAGGAGCTTTGGAGAAACAAATAAGGGCAGT 1291
Qy 1021 TGTGACGGCTCTCGCAAGAGAAAGAGATGCAAAACACTCCACGGCGAAGCCCTCC 1080
Db 1292 TGTGACGGCTCTCGCAAGAGAAAGAGATGCAAAACACTCCACGGCGAAGCCCTCC 1351
Qy 1081 GTGACGGGATCCCAAGTTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGACGGGATCCCAAGTTTACCCAGTAGCAGAGAGAGCCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAGAGCTTTACAGAACTACACACAGCTGTCTTGCACTCCAGGCTCCACAAAGAGAC 1200
Db 1412 GGCAGAGCTTTACAGAACTACACACAGCTGTCTTGCACTCCAGGCTCCACAGAGAGAC 1471
Qy 1201 CGGAGGGCCGGCGGAGTCCGCCACCATGTCTGTGGAAGGAGGAGCGCGGAGCGTGT 1260
Db 1472 CGGAGGGCCGGCGGAGTCCGCCACCATGTCTGTGGAAGGAGGAGCGCGGAGCGTGT 1531
Qy 1261 TCTCTGACCTCGCGGCGCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGAGTGGT 1320
Db 1532 TCTCTGACCTCGCGGCGCTCTGGATGAAATCGAGCCGTGGATCGAGGGGAGGAGTGGT 1591
Qy 1321 TCTGAGAGCGGATCTGAGGATGGCTTCCCAAGAGATCCATCTGGATGAAATGATGAT 1380
Db 1592 TCTGAGAGCGGATCTGAGGATGGCTTCCCAAGAGATCCATCTGGATGAAATGATGAT 1651
Qy 1381 GGAGGAAAAATPAAACATCTTACATCTTCAAGAGAGTGTAGTATTGTGGAAGTTTTTC 1440
Db 1652 GGAGGAAAAATPAAACATCTTACATCTTCAAGAGAGTGTAGTATTGTGGAAGTTTTTC 1711
Qy 1441 CGTTCAAATTTATCTCAATATTCACTGAGACGCATACAGGTGAAAAACCATACAAA 1500
Db 1712 CGTTCAAATTTATCTCAATATTCACTGAGACGCATACAGGTGAAAAACCATACAAA 1771
Qy 1501 TGTGAATTTTGTGAATATGTCGACCCAGAGACATCTCTGAGGTATCACTTTGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATATGTCGACCCAGAGACATCTCTGAGGTATCACTTTGAGAGA 1831
Qy 1561 CATCAAGGAAAAACAAACCGATTTGCTGCTGAGGTCAAGAACGATGGTAAATAATCAG 1620
Db 1832 CATCAAGGAAAAACAAACCGATTTGCTGCTGAGGTCAAGAACGATGGTAAATAATCAG 1891
Qy 1621 GACACTGAGATGACATTTAAACCGCTGACAGTGGGCAACCAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAGATGACATTTAAACCGCTGACAGTGGGCAACCAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTGCAAAAGATGTTTACAGGAGTCCACTCTGCAAGAGAGCTTTAAGAGATGCT 1740
Db 1952 TTTGATGTGCAAAAGATGTTTACAGGAGTCCACTCTGCAAGAGAGCTTTAAGAGATGCT 2011
Qy 1741 TCTGTTTTTTCAGAAATGTTTCTGGGCGAGCGCTGTCTCTCACCAGCACACAAAGATATCAG 1800

PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227709P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPT; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
PS Disclosure; SEQ ID NO 35905; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/hematopoietic antigen genomic

CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention
XX Sequence 14906 BP; 3959 A; 3088 C; 3333 G; 4526 T; 0 U; 0 Other;
SQ Query Match 49.0%; Score 1561.2; DB 4; Length 14906;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1482 AGGTGAAACCAATCAATGTAATTTTGTGAATGCTGAGCCCAAGACATCTCT 1541
DB AGGTGAAACCAATCAATGTAATTTTGTGAATGCTGAGCCCAAGACATCTCT 4748
QY 1542 GAGGTATCACTGGAGAGACATCAAGGAAACCAACCGATGCTGCTGAAGTCAA 1601
DB GAGGTATCACTGGAGAGACATCAAGGAAACCAACCGATGCTGCTGAAGTCAA 4808
QY 1602 GAACGATGTAATAATCAAGACATCAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 1661
DB GAACGATGTAATAATCAAGACATCAAGATGCACTATTAAACCGCTGACAGTGCACAAAC 4868
QY 1662 CAAATAATTTGAAAGATTTTGTGATGCTGCAAGATGTTACAGGAGTCCACTTGCAAA 1721
DB CAAATAATTTGAAAGATTTTGTGATGCTGCAAGATGTTACAGGAGTCCACTTGCAAA 4928
QY 1722 GCAGCTTAAGGAGATGCTTCTGTTTTCAGATGTTCTGGCAGGCTGCTCTCACC 1781
DB GCAGCTTAAGGAGATGCTTCTGTTTTCAGATGTTCTGGCAGGCTGCTCTCACC 4988
QY 1782 AGCACAAAGATATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
DB AGCACAAAGATATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 5048
QY 1842 GAATAAAACCCCTACCCCTGCTTACCTGGAACCTGTTAAATAAGAGATCAGCAATTTGAAAC 1901
DB GAATAAAACCCCTACCCCTGCTTACCTGGAACCTGTTAAATAAGAGATCAGCAATTTGAAAC 5108
QY 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 1961
DB TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 5168
QY 1962 TACCACCATTAACCTTGAAGTTAGCCCAAGAGAGCAACCGAGACCGCAGCTGACTG 2021
DB TACCACCATTAACCTTGAAGTTAGCCCAAGAGAGCAACCGAGACCGCAGCTGACTG 5228
QY 2022 CAGATACAGGCCAAGTGTGATGTCACGAAACCTTTAAATTTATCCGTGGGGCTCT 2081
DB CAGATACAGGCCAAGTGTGATGTCACGAAACCTTTAAATTTATCCGTGGGGCTCT 5288
QY 2082 TCACAAATTTGCCGGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATTT 2141
DB TCACAAATTTGCCGGCAATTTCTTTGAGTAAAGTTGATTCAGATATCACTGTCCATTT 5348
QY 2142 TTGTACCTTCAAGACATTTATCCAGAGATTTTAAATGATGACCAAGAGACTGGACATAA 2201
DB TTGTACCTTCAAGACATTTATCCAGAGATTTTAAATGATGACCAAGAGACTGGACATAA 5408
QY 2202 ATACAATCTGACGTTTCATAAAACCTGCGAAACCAAGTCTTTGCTTAGAAGTGCAGTAC 2261
DB ATACAATCTGACGTTTCATAAAACCTGCGAAACCAAGTCTTTGCTTAGAAGTGCAGTAC 5468
QY 2262 CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
DB CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 5528
QY 2322 CAAAGCCCAAGTCTGTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAGGGGAAGCA 2381
DB CAAAGCCCAAGTCTGTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAGGGGAAGCA 5588
QY 2382 GAGCCCTCTCGGCCAGGCAAGGCCCTCTGACTTTTCAAGGATAGACTCTAGCACTTTAGC 2441
DB GAGCCCTCTCGGCCAGGCAAGGCCCTCTGACTTTTCAAGGATAGACTCTAGCACTTTAGC 5648

QY 2442 CCCAGTAACCTGAAGTCCCAAGACCCACAGCAGAAATGTGGGGTCCCAAGGGCCGCCAC 2501
DB CCCAGTAACCTGAAGTCCCAAGACCCACAGCAGAAATGTGGGGTCCCAAGGGCCGCCAC 5708
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAAGAC 2561
DB CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAAGAC 5768
QY 2562 AAAAGACCCGAGACAAAATTTGAAACCTTCTCAGTAGTCTCTTCTCAGCCCACTCCCG 2621
DB AAAAGACCCGAGACAAAATTTGAAACCTTCTCAGTAGTCTCTTCTCAGCCCACTCCCG 5828
QY 2622 CAGCAGTAAACATCAATGTTCCATCGACTACCCCGCCAGAACAGACAGCCCTGGGCACC 2681
DB CAGCAGTAAACATCAATGTTCCATCGACTACCCCGCCAGAACAGACAGCCCTGGGCACC 5888
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGCAGCAGAAATTTGGTGA 2741
DB TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGCAGCAGAAATTTGGTGA 5948
QY 2742 GCGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTGGCTTGAAGTTCAGCAGCCCGGGGC 2801
DB GCGCCCTTCCAAAAGACTGAAAGTCCAGCGTGGTGGCTTGAAGTTCAGCAGCCCGGGGC 6008
QY 2802 CAATTACAGAAAGAGGCTATGACCTTCCCAAGTACCATAATGTTTGGTTCAGAGGCATC 2861
DB CAATTACAGAAAGAGGCTATGACCTTCCCAAGTACCATAATGTTTGGTTCAGAGGCATC 6068
QY 2862 GTTACCGCAGGACTGTGTATCCCGCAGCGCTGCTCCCAACCAAGGTTCTCTGAG 2921
DB GTTACCGCAGGACTGTGTATCCCGCAGCGCTGCTCCCAACCAAGGTTCTCTGAG 6128
QY 2922 CTCCAGCAGGCTCGATTTCTCCAAATGTCCTGATCTTTCAGAAAGCCCTATGTTGGTCC 2981
DB CTCCAGCAGGCTCGATTTCTCCAAATGTCCTGATCTTTCAGAAAGCCCTATGTTGGTCC 6188
QY 2982 GCCATTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
DB GCCATTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6248
QY 3042 TGGTGGATGTGAGT 3055
DB TGGTGGATGTGAGT 6262

RESULT 9
ACN44986
ID ACN44986 standard; DNA; 36022 BP.
XX AC ACN44986;
XX DT 18-NOV-2004 (first entry)
XX Human genomic sequence hCG37127.
XX DE Cytostatic; carcinoma; lymphoma; cancer; human; gene; ss.
XX KW Homo sapiens.
XX OS WO2003073826-A2.
XX PN 12-SEP-2003.
XX PD 28-FEB-2003; 2003WO-US006235.
XX PF 01-MAR-2002; 2002US-00087192.
XX PR (SAGR-) SAGRES DISCOVERY.
XX PI Morris DW;
XX DR WPI; 2003-328604/31.
XX

PT Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.

PS Claim 1; SEQ ID NO 1708; Opp; English.

XX The present invention relates to novel DNA and protein sequences which
CC are associated with carcinomas. The sequences are useful for: (i) for
CC screening drug candidates; (ii) for screening of bioactive agent capable
CC of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
CC a bioactive agent capable of modulating the activity of CAP; (iv) for
CC evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
CC carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
CC carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
CC (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
CC determining Carcinoma Associated (CA) gene copy number. In addition, the
CC CA genes are useful as DNA vaccines and the CAP are useful as markers of
CC carcinoma including lymphoma. The present sequence is one such CA coding
CC sequence. Note: This patent is an equivalent to basic patent
CC US2002182586A1, for which no sequence data was published

XX Sequence 36022 BP; 9645 A; 7326 C; 8145 G; 10906 T; 0 U; 0 Other;

Query Match	49.0%;	Score 1561.2;	DB 11;	Length 36022;
Best Local Similarity	99.5%;	Pred. NO. 0;		
Matches 1566;	Conservative	0;	Mismatches 8;	Indels 0;
Gaps	0;			
QY 1482	AGGTGAAAACCATACAAATGTAATTTGTAATATGCTGCAGCCCAAGAGACATCTCT	1541		
DB 15816	AGGTGAAAACCATACAAATGTAATTTGTAATATGCTGCAGCCCAAGAGACATCTCT	15875		
QY 1542	GAGGTATCATTGGAGAGACATCAAGAGAAAAAACAACCGATGTTGCTGCTGAAGTCAA	1601		
DB 15876	GAGGTATCATTGGAGAGACATCAAGAGAAAAAACAACCGATGTTGCTGCTGAAGTCAA	15935		
QY 1602	GAACGATCGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAAAC	1661		
DB 15936	GAACGATCGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCAAAAC	15995		
QY 1662	CAAAAATTTGAAGAATTTTTGATGTTGCCAAGATGCTTACAGGAGTCCACCTGGCAA	1721		
DB 15996	CAAAAATTTGAAGAATTTTTGATGTTGCCAAGATGCTTACAGGAGTCCACCTGGCAA	16055		
QY 1722	CGAGCTTAAGAGATGCCCTTCTGTTTTTTCAGAAATGTTTGGGCGAGCGCTGCTCTCAC	1781		
DB 16056	CGAGCTTAAGAGATGCCCTTCTGTTTTTTCAGAAATGTTTGGGCGAGCGCTGCTCTCAC	16115		
QY 1782	AGCACAAAAGATCTCAGGATTTCCATAAAATGCAAGTCACTGATGACAGTGTGATAAAGT	1841		
DB 16116	AGCACAAAAGATCTCAGGATTTCCATAAAATGCAAGTCACTGATGACAGTGTGATAAAGT	16175		
QY 1842	GAATAAAACCCCTACCCCTGCTTACCTGGACCTGTTAAAAGAGATCAGCAGTTGAAAC	1901		
DB 16176	GAATAAAACCCCTACCCCTGCTTACCTGGACCTGTTAAAAGAGATCAGCAGTTGAAAC	16235		
QY 1902	TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG	1961		
DB 16236	TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG	16295		
QY 1962	TACCACCATTAACCTTTGAGTTAGCCCAAGAGAAAGCAACCGAGACCGCAGCTGACTG	2021		
DB 16296	TACCACCATTAACCTTTGAGTTAGCCCAAGAGAAAGCAACCGAGACCGCAGCTGACTG	16355		
QY 2022	CAGATACAGGCCAAGTGGATGTCAGCAAAACCTTTAAATTTATCCGTGGGGCTCT	2081		
DB 16356	CAGATACAGGCCAAGTGGATGTCAGCAAAACCTTTAAATTTATCCGTGGGGCTCT	16415		
QY 2082	TCACAAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGATTCCAAGTATCACTGTCCATT	2141		
DB 16416	TCACAAATTTGCCCGCAATTTCTTTGAGTAAAGTTTGATTCCAAGTATCACTGTCCATT	16475		
QY 2142	TGTACCTTCAAGACATTTTATFCCAGAGTTTAAATGATGACACAGAGCTGGAGCATAA	2201		
DB 16476	TGTACCTTCAAGACATTTTATFCCAGAGTTTAAATGATGACACAGAGCTGGAGCATAA	16535		

RESULT 10
ACN44985

ID ACN44985 standard; cDNA; 3016 BP.

XX ACN44985;

XX AC

XX 18-NOV-2004 (first entry)

XX Mouse mRNA sequence mC76076.

XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ds.


```
Qy 1534 ACATCTCTGAGGTATCACTTGGAGAGACATACAAAGGAAAAACAAAC---CGATGTTGCT 1590
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1555 ACATCTCTGAGGTACCACTTGGAGAGACATACAAAGACAGACGCCGGTGGATGCTGCC 1614
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1591 GCTGAAGTCAAGAACGATGGTAAATAATCAGACACACTGAAGATGCACTATTAA---CGCT 1647
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1615 GCTGAGTCCAAAAGTGAAGCGCGAGCCAGGAGCGCAGGATGCGCTACTAACGCGTCT 1674
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1648 GACAGTGGCAAAACCAAAATTTGAAAGATTTTGGATGTTGTCGCAAGATGTTACAGGC 1707
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1675 GACAGTGGCAGACCAAAATTTAAAGAGATTTCTTGATGTTGCAAAAGATGTTAAAGGA 1734
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1708 AGTCACCTGCAAAAGCAGCTTAAAGAGATGCTTCTGTTTTTTCAGAAATGTTCTGGCAGC 1767
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1735 AGCCCACTGCAAGCAGCTTAAAGAGATGCTTCTGTTTCCAGAGTGTTC----- 1786
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1768 GCTGTCTCTCACGAGCAACAAAGATACTCAGATTTTCCATAAAATGCAAGCTGATGAC 1827
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1787 ----TCTCACGAGCACACAGCAAGATACTCAGGATTTCCATAAACATGCAAGCTGAT--- 1839
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1828 AGTCTGATTAAGTGAATAAAACCCCTACCTGCTTACCTGGACCTGTTAAAGAGAGA 1887
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1840 AGTCTGAGAAAGCGAGAGAGCCCTGCCCTACTTATCTGGACATGCAAGAGAAAGA-- 1897
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1888 TCAGCAGTTGAAACTCAGGCAAAATAACCTCATCTGTAGAACCAAGCGCGATGTTACTCCT 1947
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1998 -AAGCAGGGAGCCTCAGCCAGCAGCCTCTCTGCAGACTAGAGGGGGTGGGTCTTAA 1956
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1948 CCTCGGATGCGAGTACCAACCCATTAACCTTGAAGTTAGCCCCAAAGAGAAACAAAGGAG 2007
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1957 GCAGGGGAAGCTGG-----CCATAGGGAGAAAGATGAT 1989
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2008 ACCGAGCTGACTGCAGATACAGGCCAAGTGTGGATTGTACGAAAAACCTTTAAATTTA 2067
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1990 CAGGATGCTGACTACAGACATAAGCCCGTGTCTGACTGCCAGGACAGGCCCTTTGAATCTA 2049
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2068 TCCGTGGGGGCTCTTCAAAATGCCCCGCAATTTCTTTGAGTAAAGTTTGATTCCAAGT 2127
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2050 TCCCTTGGGCGCTCCAGCCCTGTCTGCAATCTCTTTGAGCAAGTGTCTATCCCCAGC 2109
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2128 ATCACTCTCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACACAG 2187
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2110 ATTGCTGCCCTTTTGTACTTCAAGACCTTTTATCGGAAGTCTTTATGATGACACCAG 2169
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2188 AGACTGGAGCAATAATCACTGACGTTTATATAAACTGTGCAAAACAGTCTTCTGCTT 2247
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2170 AGACTTGAGCACAGGTAACAACCTGACCCGACAAAGAACGGCAGCAAGTCTGTGCTG 2229
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2248 AGAAGTGCAGCTACCGGATGCCCGCAGCTTGTCTGGGAAAAAGATGTGCTCCCTCTCT 2307
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2230 AGGAACAGGCTACCGGTTGCCCTCGGCTTGTCTGGGAAAAAGATGTCCCTCCCTGTCT 2289
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2308 AGTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCACTCT 2367
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2290 GGCCTGCAACAGCCCAAGCCCAAGACTGCTTCTCACCACTCTGAAAGTCCCTGCACTCA 2349
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2368 GCGAAGGGGAAGCAGACCCCTCTGGGCGCAGCAGCCCTCTGACTTCCAGGGATAGAC 2427
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2350 GAGAAGGCTCGCAGCGGGGCTCTCGGGGCAAGCAAGCAAGCAAGTTCAGGACACAGAC 2409
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2428 TCTAGCATTTAGCCCCCAAGTAACTGAAAGTCCCAAGACCAAGCAAGATGTGGGGGTC 2487
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2410 AACAGCATTTAGCCCCCAAGTAACTGAAAGTACACAGGTACACACCCCAAGTCTGGGGGC 2469
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2488 CAAGGGGCGCCACCAAGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTCCT 2547
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2470 ACCAG---TGCCACCAAGCAGCAGTCAAGATGTTTTCCAAAGTGGCGTCCCTGCT 2526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2548 GCACCGGATAGACAAAAGACCCGAGACAAATTTGAAACCTTCTCCAGTAGCTCTTCT 2607
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2527 GCTATGGAATAGGTGAAGAGACCTGAGCCAAACCTGAAAGTCCCTACCAAGCCTCCCGTCT 2586
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Qy 2608 CAGCCCACTCTCGCAGCAGTAAACATCAATGGTTTCCATCGACTACCCCGCAAGAACAC 2667
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2587 CAGTCCCCCTCAGCAGCTAATAATAGCAACGGTTCTGTGAGTATCCCGTGAAGGTTGAC 2646
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2668 AGCCCGTGGGCACCTCCGGGAAGAGACTATTCTGTAAATCGGAGTCCAGCAATACTGCA 2727
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2647 GGGCCATGGGCACAGCAAGGAGAGACTACTACTGCCATCGGAATTTGCGCAGTGCAGCA 2706
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2728 GCAGAAATTTGTTGAGGAGCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGACGTT 2787
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2707 GCAGAGTACAGTGAAGCACATCCCAAAAGACTCAAGTCCAGTGGGTGCTCCCTGGACACA 2766
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2788 GACCAGCCGGGGCCAAATTACAGAAGAGGCTATCACTTCCCAAGTACCATATGGTACAGA 2847
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2767 GAGCATGCAGGACCAATGCGAGAAGGGCTTTGAGCTCCCCAAAGTACCATGTGTGTCAGG 2826
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2848 GGCATCATCATCTGTTTACCGCAGACCTGTGTGATCCGTGCGAGCGCTGCCCTCCCAA 2907
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2827 AGCATCACTCTTGTCTTACCAACAGAGTGTGTGCGCCACCGCTGTCTGCCCAAAA 2886
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2908 CCAAGGTTCTGAGCTCCAGCGAGTCAATTTCTCCAAATGTGCTGACTGTTTCAGAAAGCCC 2967
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2887 GCCGTTTCTGAGCCCTGGGAGGTGAGTCAACCCAGTGTGTTGGCTGTGCGAAGCCC 2946
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2968 TATGTTGGTCCCGGCCACTTTTACATTTGTGTGCTGCTGTGTAGTCCAGCATCCAGTCTG 3027
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2947 TACAGTGCCTCTGAGACCCCTGTATACCTGTGAGCCCGTGGGACACCGCAGGAGCGACCA 3006
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3028 ACGTTAGAG 3037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 3007 GCCCTTGAAG 3016
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 11

AAV09023
ID AAV09023 standard; DNA; 10282 BP.

XX AAV09023;

XX AC AAV09023;

XX DT 21-JUL-1998 (first entry)

XX XX Homo sapiens 20q13 amplicon ZABC-1 genomic sequence.

XX DE 20q13 amplicon; chromosome 20; tumour; detection; ZABC-1 gene;
XX KW chromosomal abnormalities; probe; gene therapy; antisense inhibition;
XX KW treatment; age-related macular degeneration; retinitis pigmentataion;
XX KW Leber's congenital amaurosis; zinc finger amplified in breast cancer; ds.

XX OS Homo sapiens.

XX PN WO9802539-A1.

XX XX 22-JAN-1998.

XX PF 15-JUL-1997; 97WO-US012343.

XX PR 15-JUL-1996; 96US-00680395.

XX PR 16-OCT-1996; 96US-00731499.

XX PR 17-JAN-1997; 97US-00785532.

XX XX (REGC) UNIV CALIFORNIA.

XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;

XX WPI; 1998-110587/10.

XX DR P-PSDB; AAW23975.

XX New sequences from the 20q13 amplicon - used for detecting chromosomal

XX abnormalities, particularly tumours, and for developing products for

XX treating diseases.

XX Claim 1; Page 64-67; 91pp; English.

XX XX

CC The sequence is that of the genomic sequence of ZABC-1 (zinc finger
CC amplified in breast cancer). It maps to the core of the 20q13.2 amplicon
CC and is overexpressed in primary tumours and breast cancer cell lines
CC having 20q13.2 amplification. The exact coding region for the genomic
CC sequence is not given. The sequence can be used as a probe for the
CC detection of chromosomal abnormalities at 20q13. It and other sequences
CC isolated from the 20q13 amplicon are consistently amplified in primary
CC tumours. These sequences are useful as probes or as probe targets for
CC monitoring the relative copy number of corresponding sequences from a
CC biological sample such as tumour cells. The sequences can also be used in
CC therapeutic applications for modulating the expression of the endogenous
CC gene or the activity of the gene product. Examples of therapeutic
CC approaches include antisense inhibition of gene expression, gene therapy,
CC and monoclonal antibodies that specifically bind the gene products. The
CC products can also be used in the treatment of other diseases, e.g. age-
CC related macular degeneration, Leber's congenital amaurosis and retinitis
CC pigmentation

XX SQ Sequence 10282 BP; 2820 A; 2222 C; 2191 G; 3047 T; 0 U; 2 Other;

Query Match 46.5%; Score 1482.8; DB 2; Length 10282;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1517; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 1482 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCAGAGACATCTCT 1541
DB 8204 AGGTGAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCAGAGACATCTCT 8263
QY 1542 GAGGTATCATCTGGAGAGACATCAAGGAAAAACAAACCCGATGTGCTGAGTCAA 1601
DB 8264 GAGGTATCATCTGGAGAGACATCAAGGAAAAACAAACCCGATCTTCTGCTCAACTCA 8323
QY 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGACATATTAACCGCTGACAGTGCGCAAA 1661
DB 8324 GAACGATGGTAAAAATCAGGACACTGAAGATGACATATTAACCGCTGACACTCCGCAAA 8383
QY 1662 CAATAATTTGAAAAGATTTTGTGAGTGCCAAAGATCTTACAGGACCTGACCTGCAAA 1721
DB 8384 CAATAATTTGAAAAGATTTTGTGAGTGCCAAAGATCTTACAGGACCTGACCTGCAAA 8443
QY 1722 GCAGCTTAAGGAGATGCCTCTCTGTTTTTTCAGAAATGTTCTGGGACGCTGCTCTCACC 1781
DB 8444 CCACCTTAAGGAGATGCCTCTCTGTTTTTCAGAAATGTTCTGCCCCCCTCTCTCTCACC 8503
QY 1782 AGCACAAAAGATATCAGGATTTCCATAAAATCAGCTGATGACAGTGCTGATAAAGT 1841
DB 8504 ACCACAAAAGATATCAGGATTTCCATAAAATCAGCTGATGACAGTGCTGATAAAGT 8563
QY 1842 GAATAAAACCTTACCCCTGCTTACCTGGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1901
DB 8564 GAATAAAACCTTACCCCTGCTTACCTGGAACCTGTTAAAGAGAGATCAGCAGTTGAAAC 8623
QY 1902 TCAGGCAATTAACCTTACCTGGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 1961
DB 8624 TCAGGCAATTAACCTTACCTGGAACCAAGCGGATGTTACTCTCTCCGATGGCAG 8683
QY 1962 TACCACCATTAACCTTGAAGTTAGCCCCAAAGAGAGCAAAACCGGACCGCAGCTGACTG 2021
DB 8684 TACCACCATTAACCTTGAAGTTAGCCCCAAAGAGAGCAAAACCGCAGCTGACTG 8743
QY 2022 CAGATACAGGCCAAGTGGATGTGCAGAAAAACCTTTAAATTTATTCGTTGGGGCTCT 2081
DB 8744 CAGATACAGGCCAAGTGGATGTGCAGAAAAACCTTTAAATAAATTCGTTGGGGCTCT 8803
QY 2082 TCACAAATTTGCCGGCAATTTCTTTCAGTAAAGTTTGTATCCCAAGTATCAGCTGCCATT 2141
DB 8804 TCACAAATTTGCCGGCAATTTCTTTCAGTAAAGTTTGTATCCCAAGTATCAGCTGCCATT 8863
QY 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTATATGATGCAACAGAGACTGGAGCATAA 2201
DB 8864 TTGTACCTTCAAGACATTTTATCCCAACTTTTATATCATGCAACAGACTCCACCATAA 8923
QY 2202 ATACAACTCTGACGTTTCAATAAAACCTGCGAAACAAGTCTTGTAGAACTGACGATAC 2261

DB 8924 ATACAACTCTGACGTTTCATAAAAACTGTGAAACAAGTCTTGTCTTACACTCGACGTAC 8983
QY 2262 CGGATGCCCGCCAGCGTTGTCTGGGAAAAGATGTGCTCCCTCTCTCTAGTTCTTGTAAACC 2321
DB 8984 CCGATGCCCGCCAGCGTTGTCTGGGAAAAGATGTGCTCCCTCTCTCTAGTTCTTGTAAACC 9043
QY 2322 CAAGCCCAAGTCTGCTTTCCCGGGCAGTCAAAATCCCTGCAATCTGCAAGGGGAAGA 2381
DB 9044 CAAGCCCAAGTCTGCTTTCCCGGGCAGTCAAAATCCCTGCAATCTGCAAGGGGAAGA 9103
QY 2382 GAGCCCTCTCTGGCCAGCAGCAAGGCCCTCTGACTTTTCAGGGATAGACTCTAGCAGTTTACG 2441
DB 9104 GAGCCCTCTCTGGCCAGCAGCAAGGCCCTCTGACTTTTCAGGGATAGACTCTAGCAGTTTACG 9163
QY 2442 CCCAAGTAACCTGAAGTCCCAAGACCAAGCAAGTGTGGGGTCTCAAGGGGGCCGCCAC 2501
DB 9164 CCCAAGTAACCTGAAGTCCCAAGACCAAGCAAGTGTGGGGTCTCAAGGGGGCCGCCAC 9223
QY 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTGCTGACCGGATAAGAC 2561
DB 9224 CAGGCAACAGCAATCTGAGATGTTTCTTAAACCAAGTGTTCCTGCTGACCGGATAAGAC 9283
QY 2562 AAAAGAGCCCGAGACAAATTTGAAACCTTCTCCAGTAGTCTCTTCTCAGCCACCTCGG 2621
DB 9284 AAAAGAGCCCGAGACAAATTTGAAACCTTCTCCAGTAGTCTCTTCTCAGCCACCTCGG 9343
QY 2622 CAGCAGTAACATCAATGGTTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 2681
DB 9344 CAGCAGTAACATCAATGGTTTCCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 9403
QY 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGAGCAGCAATTTGGTGA 2741
DB 9404 TCCGGCAACAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGAGCAGCAATTTGGTGA 9463
QY 2742 GCCCTTCCAAAAGACTGAAAGTCCAGCGTGTGCTTCCCTTCAAGTGTGACAGCCCGGGGC 2801
DB 9464 GCCCTTCCAAAAGACTGAAAGTCCAGCGTGTGCTTCCCTTCAAGTGTGACAGCCCGGGGC 9523
QY 2802 CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCATCACAATCACT 2861
DB 9524 CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGTTGTCAGAGGCATCACAATCACT 9583
QY 2862 GTTACCGCAGACTGTGTGATCCGTGCGAGGCGTGCCTCCCAAAACCAAGTTTCTGAG 2921
DB 9584 GTTACCGCAGACTGTGTGATCCGTGCGAGGCGTGCCTCCCAAAACCAAGTTTCTGAG 9643
QY 2922 CTCCAGGAGGTGCATTTCTCCAAATGTGCTGACTGTTTCAGAAAGCCCTATGTTGGTCCGG 2981
DB 9644 CTCCAGGAGGTGCATTTCTCCAAATGTGCTGACTGTTTCAGAAAGCCCTATGTTGGTCCGG 9703
QY 2982 GCCACTTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
DB 9704 GCCACTTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9763
QY 3042 TGGTGGATGTCACT 3055
DB 9764 TGCATGAGGGCGT 9777

RESULT 12

AAV04698

ID AAV04698 standard; cDNA; 1507 BP.

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

XX AC AAV04698;

Homo sapiens 20q13 amplicon cc49 transcript.

20q13 amplicon; chromosome 20; tumour; detection; C2H2 zinc finger;
chromosomal abnormalities; probe; gene therapy; antisense inhibition;
treatment; age-related macular degeneration; retinitis pigmentation;

KW Leber's congenital amaurosis; ds.
XX Homo sapiens.
XX WO9802539-A1.
XX 22-JAN-1998.
XX 15-JUL-1997; 97WO-US012343.
XX 15-JUL-1996; 96US-00680395.
XX 16-OCT-1996; 96US-00731499.
XX 17-JAN-1997; 97US-00785532.
XX (REGC) UNIV CALIFORNIA.
XX Gray JW, Collins CC, Hwang S, Godfrey T, Kowbel D, Rommens J;
XX WPI; 1998-110587/10.
XX
XX New sequences from the 20q13 amplicon - used for detecting chromosomal
XX abnormalities, particularly tumours, and for developing products for
XX treating diseases.
XX
XX Claim 1; Page 61-62; 91pp; English.
XX
XX The sequence is that of a cDNA sequence cc49, which was isolated from the
XX 20q13 amplicon. It shows homology to C2H2 zinc finger genes. It can be
XX used as a probe for the detection of chromosomal abnormalities at 20q13.
XX It and other sequences isolated from the 20q13 amplicon are consistently
XX amplified in primary tumours. These sequences are useful as probes or as
XX probe targets for monitoring the relative copy number of corresponding
XX sequences from a biological sample such as tumour cells. The sequences
XX can also be used in therapeutic applications for modulating the
XX expression of the endogenous gene or the activity of the gene product.
XX Examples of therapeutic approaches include antisense inhibition of gene
XX expression, gene therapy, and monoclonal antibodies that specifically
XX bind the gene products. The products can also be used in the treatment of
XX other diseases, e.g. age-related macular degeneration, Leber's congenital
XX amaurosis and retinitis pigmentata
XX
XX Sequence 1507 BP; 445 A; 354 C; 364 G; 340 T; 0 U; 4 Other;
XX
XX Query Match 34.7%; Score 1106.2; DB 2; Length 1507;
XX Best Local Similarity 96.6%; Pred. No. 0;
XX Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;
XX
XX 1 ATGCATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGGG 60
XX 320 ATGCAATCGAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGATGG 378
XX
XX 61 CCAGAAGTGATGGCAGCTCTTGGCAGTCCGATGGAGATGGAGGATGCCCTTGTCAATG 120
XX 379 CCAAGAGTGATGGCAGCTCTTGGCAGTCCGATGGAGATGGA-GATGCCCTTGTCAATG 437
XX
XX 121 AAAGGACCGCTGTGTGTTCATTCGAGCTACACAAGAAAAAATGTC-ATCCAAATCGA 179
XX 438 AAAGGCGCCNCTGTGTCAATTCGAGCTACACAAGAAAAAATGTCAATCCGAATCGA 497
XX
XX 180 GGGGTATATGCCCTTGGATTTGATGTTCTTGAGCCAGACCTTTCACATTCAGAAGACCT 239
XX 498 GGGGAATATGCCCTTGGATTTGATGTTCTTGAGCCAGACCTTTCACATTCAGAAGACCT 557
XX
XX 240 TAATAAATGTCCTTAATGCAACCGCCCTTACCTCTGTGAACAGCAGTCTTTCGGGT 299
XX 558 TAATAAATGTCCTTAATGCAACCGCCCTTACCTCTGTGAACAGCAGTCTTTCGGGT 617
XX
XX 300 TGAAGCAGATATCTCAGTCCGCTTGATATAAAGTCAAGTGGGAAACAGAACCTCCCAAGGA 359
XX 618 TGAAGCAGATATCTCAGTCCGCTTGATATAAAGTCAAGTGGGAAACAGAACCTCCCAAGGA 677
XX
XX 360 AAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGATGCGC 419
XX
Db 678 AAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGATGCGC 737
Qy 420 TTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTTCACTTACGGGTGAACAT 479
Db 738 TTTTGATGTTGAGATCCACATGAGAACACACAAAGATTCTTTTCACTTACGGGTGAACAT 797
Qy 480 GTGCGGAAGAAGTTSSRRSSAGCCTTGCTTTTAAATAATCAATGCGGACACATTAATGG 539
Db 798 GTGCGGAAGAAGATTCAAGGAGCCTTGCTTTTAAATAATCAATGCGGACACATTAATGG 857
Qy 540 CAATCGGGGCGCAGAGCAAACTGCAGCAAGGCTTCGAGAGTAGTGCAGCAACGATCAA 599
Db 858 CAATCGGGGCGCAGAGCAAACTGCAGCAAGGCTTCGAGAGTAGTGCAGCAACGATCAA 917
Qy 600 CGAGTGTCTCAGGTGCACGCGCGGAGAGCATCTCCTCTCTTACAAAATCTGCATGGT 659
Db 918 CGAGTGTCTCAGGTGCACGCGCGGAGAGCATCTCCTCTCTTGCAAAATCTGCATGGT 977
Qy 660 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCGCAAGGTGCACACAA 719
Db 978 TTGTGGCTTCTATTTCCAAATAAAGAAAGTCTTAATTGAGCACCGCAAGGTGCACACAA 1037
Qy 720 AAAAATCTGCTTTCGTTACCGAGCGCGAGAGACTCTCCACAAGGAGAAATGCGGTC 779
Db 1038 AAAAATCTGCTTTCGTTACCGAGCGCGAGAGACTCTCCACAAGGAGAAATGCGGTC 1097
Qy 780 CTCGAGGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAAATCTCACCTCGAAACGGG 839
Db 1098 CTCGAGGAGGACTTCTCGAGTTGTTCAACTTGAGACCAAAAATCTCACCTCGAAACGGG 1157
Qy 840 GAAGAAGCTGTAGATGTCATTCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCA 899
Db 1158 GAAGAAGCTGTAGATGTCATTCCTCAGCTCGATCCGTTTCAACCTTCCAGGCTTGGCA 1217
Qy 900 GCTCGCTACCAAGGAAAGTTGCCATTTCGCAAGAGTGAAGGAATCGGGGCAAGAGG 959
Db 1218 KCTGGCTACCAAGGAAAGTTGCCATTTCGCAAGAGTGAAGGAATTCGGGCAAGAGG 1277
Qy 960 GAGCACCGACAAACGACGATTCGAGTTCGAGAAAGAGCTTGGAGAAAACAAATAAGGCAG 1019
Db 1278 GAGCACCGACAAACGACGATTCGAGTTCGAGAAAGAGCTTGGAGAAAACAAATAAGNACCA 1337
Qy 1020 TTGTGAGGCTCTCGCAAGAGAAAGAGTGAAGGAATTCGCAAGAGTGAAGGAATTCGCAAG 1079
Db 1338 TTGTGAGGCTCTCGCAAGAGAAAGAGTGAAGGAATTCGCAAGAGTGAAGGAATTCGCAAG 1397
Qy 1080 CGTGAGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTCTCCGAGTG 1139
Db 1398 CGTGAGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTCTCCGAGTG 1457
Qy 1140 CGGCAAGCTTTTCAGAACCTTACCAACCTTACCAACCTTACCAACCTTACCAACCTTAC 1189
Db 1458 CGGCAAGCTTTTCAGAACCTTACCAACCTTACCAACCTTACCAACCTTACCAACCTTAC 1507
XX
XX RESULT 13
XX ACN44984
XX ID ACN44984 standard; DNA; 26345 BP.
XX XX ACN44984;
XX AC ACN44984;
XX XX 18-NOV-2004 (first entry)
XX XX Mouse genomic sequence MCG6549.
XX XX Cytostatic; carcinoma; lymphoma; cancer; murine; gene; ss.
XX XX Mus musculus.
XX XX WO2003073826-A2.
XX XX 12-SEP-2003.
XX XX

PF 28-FEB-2003; 2003WO-US006235.
PR 01-MAR-2002; 2002US-00087192.
XX (SAGR-) SAGRES DISCOVERY.
XX Morris DW;
XX WPI; 2003-328604/31.
XX Recombinant nucleic acid useful for diagnosis and treatment of carcinoma
XX comprises a nucleotide sequence.
XX Claim 1; SEQ ID NO 1705; Opp; English.
XX The present invention relates to novel DNA and protein sequences which
XX are associated with carcinomas. The sequences are useful for: (i) for
XX screening drug candidates; (ii) for screening of bioactive agent capable
XX of binding to Carcinoma Associated Protein (CAP); (iii) for screening of
XX a bioactive agent capable of modulating the activity of CAP; (iv) for
XX evaluating the effect of a candidate carcinoma drug; (v) for diagnosing
XX carcinoma; (vi) for inhibiting the activity of CAP; (vii) for treating
XX carcinoma; (viii) for neutralizing the effect of CAP; (ix) as a biochip;
XX (x) for diagnosing carcinoma or a propensity to carcinoma; and (xi) for
XX determining Carcinoma Associated (CA) gene copy number. In addition, the
XX CA genes are useful as DNA vaccines and the CAP are useful as markers of
XX carcinoma including lymphoma. The present sequence is one such CA coding
XX sequence. Note: This patent is an equivalent to basic patent
XX US2002182586A1, for which no sequence data was published
SQ Sequence 26345 BP; 6286 A; 5855 C; 6718 G; 7486 T; 0 U; 0 Other;
Query Match 23.0%; Score 733; DB 11; Length 26345;
Best Local Similarity 70.2%; Pred. No. 6.8e-208;
Matches 1106; Conservative 0; Mismatches 415; Indels 54; Gaps 7;
QY 1482 AGGTGAAACCATACAAATGTAATTTTGTGAATATCTGCGAGCCAGAGACATCTCT 1541
DB 14832 AGGTGAAACCATACAAATGTAATTTTGTGAATATCTGCGAGCCAGAGACATCTCT 14891
QY 1542 GAGGTATCACTTGGAGAGACATCAAGGAAAAACAAAC---CGATGTTGCTGTAAGT 1598
DB 14892 GAGGTACACTTGGAGAGACATCAAGAGACAGCAGCGGTGGATGTCGCGCTGATC 14951
QY 1599 CAAGAACGATGTAATAAATCAGGACACTGAAGATGCACTATTAA---CCGTGACAGTGC 1655
DB 14952 CAAAAGTGAAGCCGAGCCAGGAGCCGAGGATCGCTACTTAACGGCTGCTGACAGTGC 15011
QY 1656 GCNAACCAAAATTTGAAAGATTTTGTGATGTTGTCGCAAGATGTTACAGGAGTCCACC 1715
DB 15012 GCAGACCAAAATTTTAAAGAGATTTCTTTGATGGTCCCAAGATGTTAAGGGAAGCCACC 15071
QY 1716 TGCAAGCAGCTTAAGGAGATGCTCTCTGTTTTCAGAAATGTTCTGGGCGAGCTGTCT 1775
DB 15072 TGCCAAAGCAGCTTAAGGAGATGCTCTCTGTTTTCAGAGTGTTC-----TCTC 15119
QY 1776 CTCACGACACACAAGATATCTAGATTTTCATAAAATCAGCTGATGACAGTGTGA 1835
DB 15120 ACCAGCACACAGCAACGATCTACGATTTTCCATAAATCAGCTGAT---AGTGCTGA 15176
QY 1836 TAAAGTGAATTAACACCTTACCTGCTTACTGACCTGTTTAAAGAGATCAGCAGT 1895
DB 15177 GAAAGCGAGGAAGAGCCCTGCCCCCTACTTATCTGGACATGACAGAAAGAA---GCAGG 15233
QY 1896 TGAAACTCAGGCAATAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGGA 1955
DB 15234 GGAGCTCAGGCCAGCAGCCCTGCTGCGACTAGAGGGGTTGGTCTTTAGCACGGGA 15293
QY 1956 TGGCAGTACCCACCACTTGAAGTTAGCCCCAAGAGAGCAACCGAGAGCCGAGC 2015
DB 15294 AGCTGG-----CCATAGGAGAGAGATGGATCAGATGC 15326
QY 2016 TGACTGCAGATACAGGCCAAGTGGATGTTGTCACGAAAAACCTTTAAATTTATCCGTGG 2075

DB 15327 TGACTCAGACATAGCCCGGTGCTGACTGCCAGAGCAGGCTTTGAATCTATCCCTTGG 15386
QY 2076 GGCTCTTCCAAATTTGCCCGCAATTTCTTTGAGTAAAAAGTTTGAATCCCAAGTATCACCTG 2135
DB 15387 GCGCTCCACGCTGCTGCTGCAATCTCTTTGAGCAAGTGTCTGATCCCCAGCATTTGCCCTG 15446
QY 2136 TCCATTTTGTACCTTCAAGACATTTTATCAGAAAGTTTATATGATGACACAGAGACTGA 2195
DB 15447 CCCCCTTTTGTACTTTTCAAGACCTTTTATCCGGAAGTCTTATGATGACACAGAGACTTGA 15506
QY 2196 GCATAAATACAACTCTGACGTTTCAATAAACTGTGAAACCAAGTCTTGTCTTGAAGTGC 2255
DB 15507 GCACAGGTACAACTCTGACCCGCAAGAACGCGACAGCAGAGTCTGTCTGAGGAACAG 15566
QY 2256 ACGTACCCGATGCCCGCCAGCGTTTGTGGGAAAAAGATGCTCTCCCTCTCTTAGTTTCTG 2315
DB 15567 GCGTACCGGGTGCCTCCGGCTTGTCTGGGAAAAAGATGCTCTCCCTGTCTGGCCTGCA 15626
QY 2316 TAAACCCAAAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGCGCATCTGCCAAGGG 2375
DB 15627 CAAGCCCAAGGCCAAGACTGCTCTTCTACACACACTCGAAGTCCCTGCACTCAGAGAAGGC 15686
QY 2376 GAAGCAGAGCCCTCTCTGGGCCAGGCAAGGCCCTCTGACTTTCAGGATAGACTCTAGCAC 2435
DB 15687 TCGGAGGGGGCTCTGGGGCCCAAGCAAGACCCCAAGACTTCAGGACCCAGACAAACAGCAC 15746
QY 2436 TTTAGCCCCAAGTAACCTGTAAGTCCCAAGACCCAGCAGAAATGTGGGGGTCTCAAGGGGC 2495
DB 15747 TTTAGCCCCAAGTAACCTGTAAGTCCACAGGTCCACCAACCAATGCTGGGGGCCACCAG--- 15803
QY 2496 CGCCACACAGGCAAGCAATCTGAGATGTTTCTTAAACACAGTGTTCCTCTGACCCGGA 2555
DB 15804 TGCCACACAGGCAAGCAGTGTGTTTCTCCAAAGTGGCGCTCTGCTGTCTATGGA 15863
QY 2556 TAAGCAAAAAGAGCCCGAGACAAAATTTGAACCTCTTCCAGTGTCTCTCTCAGCCAC 2615
DB 15864 TAAGTGAAGAGACTGAGCCAAAACCTGAAGTCCCTACAGCTCTCCCTGTCTAGTCCCT 15923
QY 2616 CCTCGGACGAGTAACATCAATGTTTCCATCGACTACCCCGCCCAAGAACGACAGCCCGTG 2675
DB 15924 CCTCAGCAGTAATTAATAGCAAGTCTGTTGATGATATCCCGTGAAGGTTGACGGCCATG 15983
QY 2676 GGCACTCTCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATATCTGACAGCAAAAT 2735
DB 15984 GGCAAGCAAGGAGAGACTACTACTGCAATCGGAATTTCTGGCAGTGCAGCAGAGATA 16043
QY 2736 TGGTGAGCCCTTCCAAAAGACTGAAGTCCAGCGTGTGCTCCCTTGAAGTTGACCCAGCC 2795
DB 16044 CAGTGAGCCATCCCAAAAGACTCAAGTCCAGTGCAGTGTCTCTTGGACAACAGAGCATGC 16103
QY 2796 CGGGGCCAAATTAAGAAAGAGGCTATGACCTTTCCCAAGTACCATATGTTGTCAGAGGCATCAC 2855
DB 16104 AGGACCAATGGCAGAGGGGCTTTGAGCTCCCAAGTACCATGTTGTCAGAGGATCAC 16163
QY 2856 ATCACTGTTACCGCAGGACTGTGTGATTCCTGCGAGGCGTGTCTTCCAAACCAAGTT 2915
DB 16164 CTCTTGTACTACCAAGAGTGTGTGCGCCACCGCTGTGTGCTGCCCAACAAAGCCGTTT 16223
QY 2916 CCTGAGCTCCAGCAGGCTGATTTCTCCAAATGTGCTGATCTTTCAGAGCCCTATGGTGG 2975
DB 16224 CCTGAGCCCTGGGAGGTGGAGTCAACCAAGTGTGTGTGCTGTGAGAGCCCTTACAGTGC 16283
QY 2976 CTCCGGGCCCTTTTACACTTGTGTGCTGTGCTGTGTTGTCAGCATCCAGCTCGACTTGA 3035
DB 16284 CTCTGGACCTCTGTATACCTGTGTGACCCGTTGGACACCGCAGGAGCCAGCCCTTGA 16343
QY 3036 AGGTCTTGGTGGATG 3050
DB 16344 AGGTATTACTGTAGG 16358

RESULT 14

AAK81095
ID AAK81095 standard; DNA; 566 BP.
XX
AC AAK81095;
XX
DT 07-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:35907.
XX
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.

PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Disclosure; SEQ ID NO 35907; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
XX Sequence 566 BP; 170 A; 127 C; 134 G; 135 T; 0 U; 0 Other;
XX
XX Query Match 17.7%; Score 563.6; DB 4; Length 566;
XX Best Local Similarity 98.9%; Pred. No. 4.7e-158;
XX Matches 560; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
XX
QY 113 TGTCAATGAAGGCGCGCTGTGTTCATTCCAGCTACACAGAAAAAATGTCATCC 172
DB 1 TGTCAATGAAGGCGCGCTGTGTTCATTCCAGCTACACAGAAAAAATGTCATCC 60
XX
QY 173 AAATCGAGGGTATATGCCCTTGGATTCGATGTTCTGACGCGAGACCTTCACACATTCAG 232
DB 61 AAATCGAGGGTATATGCCCTTGGATTCGATGTTCTGACGCGAGACCTTCACACATTCAG 120
XX
QY 233 AAGACCTTAATAACATGCTTAATGCAACACCGGCTACCTCTGTGAACACGAGTTC 292
DB 121 AAGACCTTAATAACATGCTTAATGCAACACCGGCTACCTCTGTGAACACGAGTTC 180
XX
QY 293 TTCGGGTTGAACGAGATATCTCAGTCGCTTGTATAAAGTCAAGTGGGACAGACCTTC 352
DB 181 TTCGGGTTGAACGAGATATCTCAGTCGCTTGTATAAAGTCAAGTGGGACAGACCTTC 240
XX
QY 353 CCAAGGAAAGAAATTCGAGGAAATGAATTTAGCTGTGAGTATGTGGGACAGATTTA 412
DB 241 CCAAGGAAAGAAATTCGAGGAAATGAATTTAGCTGTGAGTATGTGGGACAGATTTA 300
XX
QY 413 GAGTCGCTTTTGGATGTCATGAGATCCATGAGAAACACACAAAGATTTCTTACCTACGGGT 472
DB 301 GAGTCGCTTTTGGATGTCATGAGATCCATGAGAAACACACAAAGATTTCTTACCTACGGGT 360
XX
QY 473 GTAACATGTGCGGAAGAGATTTSSAGCCCTTGTCTTAAATACATGATGCGGACAC 532
DB 361 GTAACATGTGCGGAAGAGATTTCAAGGAGCCCTTGTCTTAAATACATGATGCGGACAC 420
XX

QY 533 ATAATGGCAAAATCGGGGCGCAGAACAACTGCAGCAAGGCTTGAGAGTAGTCTCAGCAA 592
DB 421 ATAATGGCAAAATCGGGGCGCAGAACAACTGCAGCAAGGCTTGAGAGTAGTCTCAGCAA 480
XX
QY 593 CGATCAACGAGGTGCTCCAGGTGCACGGCGGAGAGCATCTCTCTCTTACAAATCT 652
DB 481 CGATCAACGAGGTGCTCCAGGTGCACGGCGGAGAGCATCTCTCTCTTACAAATCT 540
XX
QY 653 GCATGTTTGTGGCTTCTCTTATTTCCA 678
DB 541 GCATGTTTGTGGCTTCTCTTATTTCCA 566
XX
RESULT 15
AAV88042
ID AAV88042 standard; cDNA; 469 BP.
AC AAV88042;
XX
XX 12-FEB-1999 (first entry)
XX
XX EST clone FP233.
XX
XX Expressed sequence tag; secreted protein; haematopoiesis regulator;
XX tissue growth; activin; inhibitor; tumour invasion suppressor; EST; human;
XX chemotaxis; chemokinesis; haemostasis; gene therapy; thrombolysis;
XX receptor; ligand; anti-inflammatory; tumour inhibitor; ds.
XX
XX Homo sapiens;
XX
XX WO9845437-A2.
XX
XX 15-OCT-1998.
XX
XX 10-APR-1998; 98WO-US0006956.
XX
XX 10-APR-1997; 97US-00837312.
XX
XX (GENY) GENETICS INST INC.
XX
XX Jacobs K, McCoy JM, Lavallie ER, Racie LA, Merberg D, Treacy M;
XX Spaulding V, Agostino MJ;
XX
XX WPI; 1999-070078/06.
XX
XX New polynucleotides encoding human secreted proteins - derived from e.g.
XX human blood, kidney, foetal lung, placenta, testes, brain, ovary,
XX pituitary, retina and colon cDNA libraries.
XX
XX Claim 1; Page 263; 641pp; English.
XX
XX The present sequence represents an expressed sequence tag (EST), and is a
XX polynucleotide of the invention. The polynucleotides of the invention are
XX all secreted EST sequences isolated from a variety of human tissue
XX sources. The EST sequences and proteins encoded by them are predicted to
XX have useful biological activities which would make them suitable for
XX treating, preventing or ameliorating medical conditions in humans and
XX animals, although no supporting data is given. Suggested activities
XX include nutritional activity, immune stimulating or suppressing activity,
XX haematopoiesis regulating activity, tissue growth activity,
XX activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, receptor/ligand activity, anti-inflammatory
XX activity, cadherin/tumour invasion suppressor activity, tumour inhibition
XX activity. The EST sequences are also stated to be useful for gene therapy
XX
XX Sequence 469 BP; 125 A; 119 C; 132 G; 93 T; 0 U; 0 Other;
XX
XX Query Match 13.9%; Score 444.4; DB 2; Length 469;
XX Best Local Similarity 99.8%; Pred. No. 2.9e-122;
XX Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 1032 CTCGCAAGAGAAAGAGAGTGCACAACTCCACGCGGAGCGCCCTCCGTGGACGCGA 1091

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 08:43:42 ; Search time 502.445 Seconds
(without alignments)
10375.608 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCAATCGAAGTACAGG.....GGAACACTACAGTTGTGTAA 3186

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3183.6	99.9	3186	4	US-08-892-695-10 Sequence 10, Appl
2	3040.6	95.4	5632	3	US-09-560-594-3 Sequence 3, Appl
3	3040.6	95.4	5632	4	US-09-949-016-862 Sequence 862, App
4	3040.6	95.4	5632	4	US-09-949-016-862 Sequence 4262, Ap
5	1561.2	49.0	20022	4	US-09-949-016-12604 Sequence 12604, A
6	1561.2	49.0	20023	4	US-09-949-016-16004 Sequence 16004, A
7	1559.6	49.0	10365	4	US-08-892-695-9 Sequence 9, Appl
8	1106.2	34.7	1507	2	US-08-680-395-3 Sequence 3, Appl
9	1106.2	34.7	1507	4	US-08-892-695-3 Sequence 3, Appl
10	600.6	18.9	601	4	US-09-949-016-34700 Sequence 34700, A
11	600.6	18.9	601	4	US-09-949-016-151880 Sequence 151880, A
12	598.2	18.8	601	4	US-09-949-016-34699 Sequence 34699, A
13	598.2	18.8	601	4	US-09-949-016-151879 Sequence 151879, A
14	66.8	2.1	2765	4	US-09-620-312D-61 Sequence 61, Appl
15	66.8	2.1	2799	4	US-09-949-016-4961 Sequence 4961, Ap
16	66.8	2.1	3039	4	US-09-949-016-1038 Sequence 1038, Ap
17	66.8	2.1	12323	4	US-09-949-016-16703 Sequence 16703, A
18	59.8	1.9	2785	4	US-09-774-528-105 Sequence 105, App
19	59.2	1.9	696	4	US-09-451-651-4 Sequence 4, Appl
20	59.2	1.9	2920	4	US-09-620-312D-1084 Sequence 1084, Ap
21	57.6	1.8	2752	4	US-09-949-016-3563 Sequence 3563, Ap
22	57.6	1.8	18824	4	US-09-949-016-15305 Sequence 15305, A
23	56.8	1.8	936	4	US-09-016-434-312 Sequence 312, App
24	55.2	1.7	2771	4	US-09-976-594-691 Sequence 691, App
25	55	1.7	4272	4	US-09-620-312D-586 Sequence 586, App
26	54.8	1.7	265	4	US-09-016-434-836 Sequence 836, App
27	54.4	1.7	2489	4	US-09-774-528-280 Sequence 280, App

ALIGNMENTS

RESULT 1

US-08-892-695-10
; Sequence 10, Application US/08892695A
; Patent No. 6808878

GENERAL INFORMATION:
; APPLICANT: Gray, Joe W

; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In

; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David

; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES

; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892.695A

; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532

; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499

; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395

; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 10

; LENGTH: 3186

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: ZABCI Open

; OTHER INFORMATION: Reading Frame

US-08-892-695-10

Query Match 99.9%; Score 3183.6; DB 4; Length 3186;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTCTTAATCTACATGATGGG 60

DB 1 ATGCAATCGAAGTACAGGAAACATGCCAACTCAATCCCTCTTAATCTACATGATGGG 60

QY 61 CCAGAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCGCTTGTCAATG 120

DB 61 CCAGAGTGTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGCGCTTGTCAATG 120

QY 121 AAAGGACCGCTGTTGTTCCATTCCAGCTACACAAGAAAAAATGTTCATCAATCGAG 180

DB 121 AAAGGACCGCTGTTGTTCCATTCCAGCTACACAAGAAAAAATGTTCATCAATCGAG 180

QY 181 GGGTATATGCCCTTGGATTGATGTTCTGCGAGCCGACCTTCACACATTCAGAGACCTT 240

DB 181 GGGTATATGCCCTTGGATTGATGTTCTGCGAGCCGACCTTCACACATTCAGAGACCTT 240

Db 181 GGGTATATGCCCCCTTGGATTGCGATGTTCTGCGAGCCAGACCTTTCACACATTCAGAAAGACCTT 240
Qy 241 AATAAAATGCTCTTAATCAACACCGGCTTACCCTCTGTGAACACGAGTCTTTCGGGTT 300
Db 241 AATAAAATGCTCTTAATCAACACCGGCTTACCCTCTGTGAACACGAGTCTTTCGGGTT 300
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGACCTTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAGTGCAGAACAGACCTTCCCAAGGAA 360
Qy 361 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGCCT 420
Db 361 AAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTGCCT 420
Qy 421 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Db 421 TTTGATGTTGAGATCCACATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Qy 481 TCGGAAGAAAGTTSRRSSAGCCTTGGTTCCTTAAAAATCACATCGCGACACATAATGGC 540
Db 481 TCGGAAGAAAGTTCAGAGGCTTGGTTCCTTAAAAATCACATCGCGACACATAATGGC 540
Qy 541 AAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGGAGAGTAGTTCAGCAACGATCAAC 600
Db 541 AAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGGAGAGTAGTTCAGCAACGATCAAC 600
Qy 601 GAGTCTGTCAGGTGCGCGCGGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 660
Db 601 GAGTCTGTCAGGTGCGCGCGGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 660
Qy 661 TGTGGCTTCCTATTTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 720
Db 661 TGTGGCTTCCTATTTTCCAAATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 720
Qy 721 AAAATGCTCTTCGGTACAGAGCGCGCAGACAGACTCTCCACAGAGGAGTAATGCGGTCC 780
Db 721 AAAATGCTCTTCGGTACAGAGCGCGCAGACAGACTCTCCACAGAGGAGTAATGCGGTCC 780
Qy 781 TCGAGGAGGACTTCTCGAGTGTTCACCTTGAGACCAAAATCTCACCCCTGAAACGSGG 840
Db 781 TCGAGGAGGACTTCTCGAGTGTTCACCTTGAGACCAAAATCTCACCCCTGAAACGSGG 840
Qy 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTACCACCTTCCAGGCTTGGCAG 900
Db 841 AAGAAGCCTGTGAGATGATCCCTCAGCTCGATCCGTTTACCACCTTCCAGGCTTGGCAG 900
Qy 901 CTGGCTACCAAGGAAAAATGTCGCTATGTCGCAAGAGTGAAGGAATCGGGGCAAGAGGG 960
Db 901 CTGGCTACCAAGGAAAAATGTCGCTATGTCGCAAGAGTGAAGGAATCGGGGCAAGAGGG 960
Qy 961 AGCACCGACACGAGATTTCGAGTTCGAGAGGAGCTTGAGAAACAAATAAGGGCAGT 1020
Db 961 AGCACCGACACGAGATTTCGAGTTCGAGAGGAGCTTGAGAAACAAATAAGGGCAGT 1020
Qy 1021 TGTGAGGCGCTCTCGCAAGAGAAAGAGTGCACCACTCCCAAGGAGCGCCCTCC 1080
Db 1021 TGTGAGGCGCTCTCGCAAGAGAAAGAGTGCACCACTCCCAAGGAGCGCCCTCC 1080
Qy 1081 GTGGAACGCGGATCCCAAGTACCCAGTAGCAGAGAGAGGCCACTCACTGCTCCGAGTGC 1140
Db 1081 GTGGAACGCGGATCCCAAGTACCCAGTAGCAGAGAGAGGCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAGCTTTCAGAACCTTACACAGCTGCTTGCATCCAGGGTCCACAGAGGAC 1200
Db 1141 GGCAAGCTTTCAGAACCTTACACAGCTGCTTGCATCCAGGGTCCACAGAGGAC 1200
Qy 1201 CGGAGGCGGCGGAGTTCGCGCCACCATGTCTGTGAGCGGAGGAGCGCGGAGCGTGT 1260
Db 1201 CGGAGGCGGCGGAGTTCGCGCCACCATGTCTGTGAGCGGAGGAGCGCGGAGCGTGT 1260
Qy 1261 TCTCTGACCTCGCGCCCTCTGATGAAATGAGCGCGTGGATCGAGGGGAGGTGGT 1320
Db 1261 TCTCTGACCTCGCGCCCTCTGATGAAATGAGCGCGTGGATCGAGGGGAGGTGGT 1320

Qy 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAAAGGAATCCATCTGGATAAAAAATGATGAT 1380
Db 1321 TCTGAAGACGATCTGAGGATGGCTTCCGAAAGGAATCCATCTGGATAAAAAATGATGAT 1380
Qy 1381 GGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTC 1440
Db 1381 GGAGGAAAAATAAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTTC 1440
Qy 1441 CGTTCAAAATTTACCTCAATATTTCACTCAGAACGCATACAGGTGAAAAACCATACAAA 1500
Db 1441 CGTTCAAAATTTACCTCAATATTTCACTCAGAACGCATACAGGTGAAAAACCATACAAA 1500
Qy 1501 TGTGAAATTTTGTGAATATGCTGCAGCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1501 TGTGAAATTTTGTGAATATGCTGCAGCCGAGAACGACATCTCTGAGGTATCACTTGGAGAGA 1560
Qy 1561 CATCACAAGGAAAAAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAAATCAG 1620
Db 1561 CATCACAAGGAAAAAAACCGATGTTGCTGTGAAGTCAAGAACGATGGTAAAAATCAG 1620
Qy 1621 GACACTGAAGTGCACATTTAAACCGCTGACAGTGCAGAACCAAAATTTGAAAAGATTT 1680
Db 1621 GACACTGAAGTGCACATTTAAACCGCTGACAGTGCAGAACCAAAATTTGAAAAGATTT 1680
Qy 1681 TTTGATGGTCCAAAGATGTTTACAGGCACTCCACTGCAAGCAGCTTAAAGGAGATGCCT 1740
Db 1681 TTTGATGGTCCAAAGATGTTTACAGGCACTCCACTGCAAGCAGCTTAAAGGAGATGCCT 1740
Qy 1741 TCTGTTTTTCAAGATGTTCTGGGCGAGCTGCTCTCAACGACACAAAGATCTCAG 1800
Db 1741 TCTGTTTTTCAAGATGTTCTGGGCGAGCTGCTCTCAACGACACAAAGATCTCAG 1800
Qy 1801 GATTTCCATAAAATGAGTGCATGACAGTGCATAAAGTGAATAAAACCCCTACCCCT 1860
Db 1801 GATTTCCATAAAATGAGTGCATGACAGTGCATAAAGTGAATAAAACCCCTACCCCT 1860
Qy 1861 GCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAA 1920
Db 1861 GCTTACCTGGAACCTGTTTAAAAAAGAGATCAGCAGTGTGAAACTCAGGCAAAATAA 1920
Qy 1921 TGTAGAACCAAGGCGGATGTTATCTCTCCGATGGCAGTACACCCATAACTTGAA 1980
Db 1921 TGTAGAACCAAGGCGGATGTTATCTCTCCGATGGCAGTACACCCATAACTTGAA 1980
Qy 1981 GTTAGCCCCCAAGAGAGCAAAACCGAGACCGCAGCTGACATGACAGTACAGGCCAAGTGTG 2040
Db 1981 GTTAGCCCCCAAGAGAGCAAAACCGAGACCGCAGCTGACATGACAGGCCAAGTGTG 2040
Qy 2041 GATTTGTCAAGAAAAACCTTTTAAATTTTATCCGTTGGGGGCTCTTTCACAAATTC 2100
Db 2041 GATTTGTCAAGAAAAACCTTTTAAATTTTATCCGTTGGGGGCTCTTTCACAAATTC 2100
Qy 2101 TCTTTGAGTAAAGTGTGATTCGAAAGTATCACTGTGCTCACTTTGTACCTTCAAGACA 2160
Db 2101 TCTTTGAGTAAAGTGTGATTCGAAAGTATCACTGTGCTCACTTTGTACCTTCAAGACA 2160
Qy 2161 TATCCAGAGTTTTTAATGATGCACAGAGACTGGGACATAAATCAATCTCTGAGTTCAT 2220
Db 2161 TATCCAGAGTTTTTAATGATGCACAGAGACTGGGACATAAATCAATCTCTGAGTTCAT 2220
Qy 2221 AAAAATGTGCAAAACAAAGTCTTGTAGAAAGTCCGATCCGGATGCCCGCCAGCGTGTG 2280
Db 2221 AAAAATGTGCAAAACAAAGTCTTGTAGAAAGTCCGATCCGGATGCCCGCCAGCGTGTG 2280
Qy 2281 CTGGAAGAGATGTGCGCTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTTC 2340
Db 2281 CTGGAAGAGATGTGCGCTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTTC 2340
Qy 2341 CCGGCGAGTCCAAATCCCTGCCATCTGCGAAGGGAGAGAGCCCTCTCTGGGCGAGGC 2400
Db 2341 CCGGCGAGTCCAAATCCCTGCCATCTGCGAAGGGAGAGAGCCCTCTCTGGGCGAGGC 2400

Qy	2401	AAGGCCCTCTGACTTCAGGGATAGACTCTAGACATTTAGCCCCAAGTAACTCGAAGTCC	2460
Db	2401	AAGGCCCTCTGACTTCAGGGATAGACTCTAGACATTTAGCCCCAAGTAACTCGAAGTCC	2460
Qy	2461	CACAGACACAGCAGAATGTGGGGTCCAAAGGGCCGCCACAGGCAACAGCAATCTGAG	2520
Db	2461	CACAGACCAAGCAGAATGTGGGGTCCAAAGGGCCGCCACAGGCAACAGCAATCTGAG	2520
Qy	2521	ATGTTTCTTAAACACGAGTGTTCCTCTGCACCGGATAAGAACAATAAGACCCGAGACAAA	2580
Db	2521	ATGTTTCTTAAACACGAGTGTTCCTCTGCACCGGATAAGAACAATAAGACCCGAGACAAA	2580
Qy	2581	TTGAAACCTCTTCAGTAGTCTCTTCACGCCACACCTCGGCAGCAGTACATCAATGGT	2640
Db	2581	TTGAAACCTCTTCAGTAGTCTCTTCACGCCACACCTCGGCAGCAGTACATCAATGGT	2640
Qy	2641	TCCATCGACTACCCCGCCAAAGAACACAGACCCGTGGGCACCTCGGGAAGAGACTATTTC	2700
Db	2641	TCCATCGACTACCCCGCCAAAGAACACAGACCCGTGGGCACCTCGGGAAGAGACTATTTC	2700
Qy	2701	TGTAAATCGGAGTGCACCAATATCTGCAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG	2760
Db	2701	TGTAAATCGGAGTGCACCAATATCTGCAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG	2760
Qy	2761	AAGTCCAGCGTGGTTGCCCTTGAGCTTGACCGCCGGGGCCAAATTACAGAGAGGCTAT	2820
Db	2761	AAGTCCAGCGTGGTTGCCCTTGAGCTTGACCGCCGGGGCCAAATTACAGAGAGGCTAT	2820
Qy	2821	GACCTTCCCAAGTACCATATGTTGAGAGGCATCACATCACTGTTTACCGCAGGACTGTGTG	2880
Db	2821	GACCTTCCCAAGTACCATATGTTGAGAGGCATCACATCACTGTTTACCGCAGGACTGTGTG	2880
Qy	2881	TATCCGTGCGAGGGCTGCCCTCCAAACCAAGGTTCTGTAGACTCCAGCGAGGTCGATTCT	2940
Db	2881	TATCCGTGCGAGGGCTGCCCTCCAAACCAAGGTTCTGTAGACTCCAGCGAGGTCGATTCT	2940
Qy	2941	CCAAATGTGCTGACTGTTTCAGAGGCCCTATGTTGGCTCCGGGCCACCTTTACACTTGTGTG	3000
Db	2941	CCAAATGTGCTGACTGTTTCAGAGGCCCTATGTTGGCTCCGGGCCACCTTTACACTTGTGTG	3000
Qy	3001	CTGTCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTTCTGGTGGTAGTCTAGTGCTTA	3060
Db	3001	CTGTCTGGTAGTCCAGCATCCAGCTCGACGTTAGAAGGTTCTGGTGGTAGTCTAGTGCTTA	3060
Qy	3061	CTCCCCATGAATTTAAATTTTACTTTCATCTCTTTGAGAGCGAATGGTGAAGCTACTGAA	3120
Db	3061	CTCCCCATGAATTTAAATTTTACTTTCATCTCTCTTTGAGAGCGAATGGTGAAGCTACTGAA	3120
Qy	3121	ATAAGCTGTGATTGTACTGTGTACATAAAACATATGAGGAATCTCGCAGAGAACCTACAGTT	3180
Db	3121	ATAAGCTGTGATTGTACTGTGTACATAAAACATATGAGGAATCTCGCAGAGAACCTACAGTT	3180
Qy	3181	GTGTAA 3186	
Db	3181	GTGTAA 3186	

```

RESULT 2
US-09-560-594-3
; Sequence 3, Application US/09560594
; Patent No. 6242590
; GENERAL INFORMATION:
; APPLICANT: Lex M. Cowseert
; TITLE OF INVENTION: ANTISENSE MODULATION OF ZINC FINGER PROTEIN-217 EXPRESSION
; FILE REFERENCE: RTS-0144
; CURRENT APPLICATION NUMBER: US/09/560,594
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 89
; SEQ ID NO 3
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```


Db 1232 AGCACCGACAAACGACGATTCGAGTTCGAGAGAGCTTGGAGAAACAAATAAGGGCAGT 1291
Qy 1021 TGTGAGGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCGGCGAAGCGCCCTCC 1080
Db 1292 TGTGAGGCGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCCGGCGAAGCGCCCTCC 1351
Qy 1081 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1352 GTGACGCGGATCCCAAGTTACCCAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGC 1411
Qy 1141 GGCAAGCTTTAGAACCTTACCAACAGCTGCTTGTGCACTCAGGCTCCACAAAGAGAC 1200
Db 1412 GGCAAGCTTTAGAACCTTACCAACAGCTGCTTGTGCACTCAGGCTCCACAAAGAGAC 1471
Qy 1201 CGGAGGCGGCGGAGTCCGCCACCATGTCGTGAGCGGAGGCGAGCGCGGAGACGTGT 1260
Db 1472 CGGAGGCGGCGGAGTCCGCCACCATGTCGTGAGCGGAGGCGAGCGCGGAGACGTGT 1531
Qy 1261 TCTCTGACCTCGCGCGCCCTCTGATGAAATGAGCGCGTGGATCCGAGGGGAAAGTGGT 1320
Db 1532 TCTCTGACCTCGCGCGCCCTCTGATGAAATGAGCGCGTGGATCCGAGGGGAAAGTGGT 1591
Qy 1321 TCTGAAGCGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATTAATAATGATGAT 1380
Db 1592 TCTGAAGCGGATCTGAGGATGGGCTTCCGAAAGGAATCCATCTGGATTAATAATGATGAT 1651
Qy 1381 GGAGGAAATTAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTC 1440
Db 1652 GGAGGAAATTAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTC 1711
Qy 1441 CGTTCAAAATTTACCTCAATATTCTCAGAACGCATACAGGTGAAAAACCATACAAA 1500
Db 1712 CGTTCAAAATTTACCTCAATATTCTCAGAACGCATACAGGTGAAAAACCATACAAA 1771
Qy 1501 TGTGAATTTTGAATATGTCGAGCCGAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Db 1772 TGTGAATTTTGAATATGTCGAGCCGAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1831
Qy 1561 CATCACAGGAAACAAACCGATGTCCTGCTGAAGTCAAGACGATGTTAAATACAG 1620
Db 1832 CATCACAGGAAACAAACCGATGTCCTGCTGAAGTCAAGACGATGTTAAATACAG 1891
Qy 1621 GACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAACCAAAAAATTTGAAAGATTT 1680
Db 1892 GACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAACCAAAAAATTTGAAAGATTT 1951
Qy 1681 TTTGATGTGCAAAAGATGTTACAGGAGTCCACTGCAAAAGCAGCTTAAAGAGATGCT 1740
Db 1952 TTTGATGTGCAAAAGATGTTACAGGAGTCCACTGCAAAAGCAGCTTAAAGAGATGCT 2011
Qy 1741 TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCCTCACCAGCACACAAAGATCTCAG 1800
Db 2012 TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCCTCACCAGCACACAAAGATCTCAG 2071
Qy 1801 GATTTCCATAAAATGCACTGATCAGCAGTGTGATAAAGTGAATAAAAAACCTACCCCT 1860
Db 2072 GATTTCCATAAAATGCACTGATCAGCAGTGTGATAAAGTGAATAAAAAACCTACCCCT 2131
Qy 1861 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTGTGAAATCTCAGGCAAAATAACCTCATC 1920
Db 2132 GCTTACCTGGACCTGTTAAAAAGAGATCAGCAGTGTGAAATCTCAGGCAAAATAACCTCATC 2191
Qy 1921 TGTAGAACCAAGCGGATGTTACTCCTCTCGGATGGCAGTACCAACCCATAACCTTGAA 1980
Db 2192 TGTAGAACCAAGCGGATGTTACTCCTCTCGGATGGCAGTACCAACCCATAACCTTGAA 2251
Qy 1981 GTTACGCCCAAGAGAACCAACCGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTG 2040
Db 2252 GTTACGCCCAAGAGAACCAACCGAGACCGCAGCTGACTGCAGATACAGGCCAAGTGTG 2311
Qy 2041 GATTTGTACGAAAAACCTTTAAATTTATCCGTTGGGGCTTTTCACAATTTGCCCGCAATT 2100

Db 2312 GATTTGTACGAAAAACCTTTAAATTTATCCGTTGGGGCTCTTTCACAAATGTCGCCGCAATT 2371
Qy 2101 TCTTTGAGTAAAAGTGTGATTTCCAAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAAGTGTGATTTCCAAGTATCACCTGTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAAGTTTTTAATGATGACACAGAGACTGGAGCATAAATACATCTCTGAAGTTCAT 2220
Db 2432 TATCCAGAAGTTTTTAATGATGACACAGAGACTGGAGCATAAATACAAATCTCTGAAGTTCAT 2491
Qy 2221 AAAAAGTGTGAAAACAGTCTTCTGTTAGAGTGCAGTACCGGATGCCGCCAGCGTTG 2280
Db 2492 AAAAAGTGTGAAAACAGTCTTCTGTTAGAGTGCAGTACCGGATGCCGCCAGCGTTG 2551
Qy 2281 CTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTTC 2340
Db 2552 CTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCCCAAGCCCAAGTCTGCTTTTC 2611
Qy 2341 CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCGCAGGC 2400
Db 2612 CCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCTGGGCGCAGGC 2671
Qy 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCAGCTTTAGCCCAAGTAACTCTGAAGTCC 2460
Db 2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCAGCTTTAGCCCAAGTAACTCTGAAGTCC 2731
Qy 2461 CACAGACACAGCAGAAATGTGGGGTCCAAGGGCGCCACACAGGCAACAGCAATCTGAG 2520
Db 2732 CACAGACACAGCAGAAATGTGGGGTCCNAGGGCGCCACACAGGCAACAGCAATCTGAG 2791
Qy 2521 ATGTTTTCTAAAAACAGTGTTCCTGCGACCCGATTAAGACAAAGAACCCGAGACAAA 2580
Db 2792 ATGTTTTCTAAAAACAGTGTTCCTGCGACCCGATTAAGACAAAGAACCCGAGACAAA 2851
Qy 2581 TTGAAAACCTCTTCAGTAGTCTCTTCAGCCGACCTCGCAGCAGTAAACATCAATGGT 2640
Db 2852 TTGAAAACCTCTTCAGTAGTCTCTTCAGCCGACCTCGCAGCAGTAAACATCAATGGT 2911
Qy 2641 TCCATCGACTACCCCGCAAGAACAGACAGCCGTCGGGACCTCCGGGAAGAGACTATTTC 2700
Db 2912 TCCATCGACTACCCCGCCNAGAACAGACAGCCGTCGGGACCTCCGGGAAGAGACTATTTC 2971
Qy 2701 TGTAAATCGGAGTCCAGCAATACCTGACAGCAATTTGTTGAGCCCTTCCAAAAAGACTG 2760
Db 2972 TGTAAATCGGAGTCCAGCAATACCTGACAGCAATTTGTTGAGCCCTTCCAAAAAGACTG 3031
Qy 2761 AAGTCCAGCGTGTGCTTGAAGTTGACAGCCCGGGCCCAATTACAGAGAGGCTAT 2820
Db 3032 AAGTCCAGCGTGTGCTTGAAGTTGACAGCCCGGGCCCAATTACAGAGAGGCTAT 3091
Qy 2821 GACCTTCCAAAGTACCATATGTTGAGGCAATCATCATGTTTACCGCAGGACTGTGTG 2880
Db 3092 GACCTTCCAAAGTACCATATGTTGAGGCAATCATCATGTTTACCGCAGGACTGTGTG 3151
Qy 2881 TATCCGTGCGAGGCGCTCCCAAAACCAAGTTCTGAGCTCCAGCGAGTTCGATTC 2940
Db 3152 TATCCGTGCGAGGCGCTCCCAAAACCAAGTTCTGAGCTCCAGCGAGTTCGATTC 3211
Qy 2941 CCAATGTGCTGCTGTTGAGAGCCCTTATGTTGCTCCGGGCCACTTTACATTTGCTGTG 3000
Db 3212 CCAATGTGCTGCTGTTGAGAGCCCTTATGTTGCTCCGGGCCACTTTACATTTGCTGTG 3271
Qy 3001 CTGCTGTGCTGAGTCCAGCATCCAGCTCCAGCTTAG----- 3034
Db 3272 CTGCTGTGCTGAGTCCAGCATCCAGCTCCAGCTTAGAGAAAAGGCTGTGTATCAATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGCGCACAAAGAGAAATATGAGAAATTTTATTTGGGAATGCACAT 3391
Qy 3035 ----- -AAGGCTCTTGGTGG 3047
Db 3392 TATCGACCAATGACAAAAAACTTTGATTCTACTAATTAGGGGGGAAAAAAGGCTCTTGGTGG 3451

QY 3048 ATGTCAGTCTTACTCCCATGAAATTAATTTTACTCTTATCTTATGAGAGCGAATGGT 3107
|||||
Db 3452 ATGTCAGTCTTACTCCCATGAAATTAATTTTACTCTTATCTTATGAGAGCGAATGGT 3511
|||||
QY 3108 GAAAGCTACTGAAATTAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3167
|||||
Db 3512 GAAAGCTACTGAAATTAAGCTGTGATTGTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
|||||
QY 3168 GAACACTACAGTTGTGTAA 3186
|||||
Db 3572 GAACACTACAGTTGTGTAA 3590
|||||

RESULT 3

US-09-949-016-862
; Sequence 862, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 862
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-862

Query Match 95.4%; Score 3040.6; DB 4; Length 5632;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

QY 1 ATGCAATCGAAGTACAGGAACATGCCAATCAATCCCTTTAATGTACATGATGGG 60
Db 272 ATGCAATCGAAGTACAGGAACATGCCAATCAATCCCTTTAATGTACATGATGGG 331
|||||
QY 61 CCAGAAATGATTGGCAGCTCTCTGGCAGTCGATGGAGATGGAGATGCCCTTTGCAATG 120
Db 332 CCAGAAATGATTGGCAGCTCTCTGGCAGTCGATGGAGATGGAGATGCCCTTTGCAATG 391
|||||
QY 121 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCTCATCCAAATCGAG 180
Db 392 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTCTCATCCAAATCGAG 451
|||||
QY 181 GGGTATATGCCCTTGGATGATGTTCTGAGCAGAGCTTCCACATTCAGAGACCTT 240
Db 452 GGGTATATGCCCTTGGATGATGTTCTGAGCAGAGCTTCCACATTCAGAGACCTT 511
|||||
QY 241 AATAAACATGCTTAATCAACACCGGCTTACCCTCTGTGAACAGAGAGTCTTCGGGTT 300
Db 512 AATAAACATGCTTAATCAACACCGGCTTACCCTCTGTGAACAGAGAGTCTTCGGGTT 571
|||||
QY 301 GAAGCAGAGATATCTCAGTCCGCTTGATTAAGTCAAGTCCGAACAGAAACCTCCCAAGGAA 360
Db 572 GAAGCAGAGATATCTCAGTCCGCTTGATTAAGTCAAGTCCGAACAGAAACCTCCCAAGGAA 631
|||||
QY 361 AAGAAATGCAAGAAAAATTAATTTAGCTGTGAGGATGTGGGCGACATTTAGAGTCGCT 420
Db 632 AAGAAATGCAAGAAAAATTAATTTAGCTGTGAGGATGTGGGCGACATTTAGAGTCGCT 691
|||||
QY 421 TTTGATGTTGAGATCCATGAGAACACACAAGATTTCTTCACTTACGGGTGTAACATG 480
|||||

Db 692 TTTGATGTTGAGATCCATGAGAACACACAAGATTTCTTTCACCTTACGGGTGTAACATG 751
QY 481 TCGCGAAGAAGMTTSRRSSAGCCTTGGTTTCTTAAAAATCAATCGGAGACATTAATGCG 540
|||||
Db 752 TCGCGAAGAAGATTCAGAGCCTTGGTTTCTTAAAAATCAATCGGAGACATTAATGCG 811
|||||
QY 541 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCCTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAAAGCAAACTGCAGCAAGCCTTGGAGAGTAGTCCAGCAACGATCAAC 871
|||||
QY 601 GAGGTCTGTCAGGTGCACGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 660
Db 872 GAGGTCTGTCAGGTGCACGCGCCGAGAGCATCTCTCTCTTCAAAAATCTGCAATGTT 931
|||||
QY 661 TGTGCTTCTTATTTTCAAAAATTAAGAAAGTCTAATGAGACCGCAAGTGCACACCAA 720
Db 932 TGTGCTTCTTATTTTCAAAAATTAAGAAAGTCTAATGAGACCGCAAGTGCACACCAA 991
|||||
QY 721 AAAACTGCTTTCGGTACCAGAGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTC 780
Db 992 AAAACTGCTTTCGGTACCAGAGCGCAGACAGACTCTCCACAAGGAGGAATGCCGTC 1051
|||||
QY 781 TCGAGGGAGGACTTCTCTGAGTTGTTCAACTTGAGACCAAAAATCTCACCTCGAAACGGG 840
Db 1052 TCGAGGGAGGACTTCTCTGAGTTGTTCAACTTGAGACCAAAAATCTCACCTCGAAACGGG 1111
|||||
QY 841 AAGAAAGCTGTGAGATGATCTCTGAGTCTGATCTGATCTGATCTGATCTGATCTGAT 900
Db 1112 AAGAAAGCTGTGAGATGATCTCTGAGTCTGATCTGATCTGATCTGATCTGATCTGAT 1171
|||||
QY 901 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAAAGTGAAGAAATCGGGGCAAGAGGG 960
Db 1172 CTGGCTACCAAGGAAAGTTGCCATTTGCCAAGAAAGTGAAGAAATCGGGGCAAGAGGG 1231
|||||
QY 961 AGCAACCAACAGCAGATTTCGAGTTCCGAGAAAGAGCTTGGAGAAACAAAATAGGCGAGT 1020
Db 1232 AGCAACCAACAGCAGATTTCGAGTTCCGAGAAAGAGCTTGGAGAAACAAAATAGGCGAGT 1291
|||||
QY 1021 TGTGAGGCTCTCTCGAAGAGAGAAAGTGCACCACTCCCAAGGAGGAGCGGCGCTTC 1080
Db 1292 TGTGAGGCTCTCTCGAAGAGAGAAAGTGCACCACTCCCAAGGAGGAGCGGCGCTTC 1351
|||||
QY 1081 GTGAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCACTCAGTCTCCAGCTGC 1140
Db 1352 GTGAGCGGATCCCAAGTTACCCAGTAGCAAGGAGAGCCACTCAGTCTCCAGCTGC 1411
|||||
QY 1141 GGCAAAGCTTTTCAAGAACTTACCACTGCTTTCGACTCCAGGCTCCAGAAAGAGAC 1200
Db 1412 GGCAAAGCTTTTCAAGAACTTACCACTGCTTTCGACTCCAGGCTCCAGAAAGAGAC 1471
|||||
QY 1201 CGGAGGCGCGCGGAGTCCGCCAATGCTGTGGAAGGAGGAGCGGCGGAGCGTGT 1260
Db 1472 CGGAGGCGCGCGGAGTCCGCCAATGCTGTGGAAGGAGGAGCGGCGGAGCGTGT 1531
|||||
QY 1261 TCTCTGACCTCGCGGCGGCTCTGGATGAAATGAGGCGGTGATCGAGGGGAGGAGTGT 1320
Db 1532 TCTCTGACCTCGCGGCGGCTCTGGATGAAATGAGGCGGTGATCGAGGGGAGGAGTGT 1591
|||||
QY 1321 TCTGAAGACGATCTGAGGATGGGCTTCCCAAGGAAATCCCATCTCGATAAAAATCATGAT 1380
Db 1592 TCTGAAGACGATCTGAGGATGGGCTTCCCAAGGAAATCCCATCTCGATAAAAATCATGAT 1651
|||||
QY 1381 GGAGAAAAATTAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTC 1440
Db 1652 GGAGAAAAATTAACATCTTACATCTTCAAGAGAGTGTAGTTATTTGGAAAGTTTTTC 1711
|||||
QY 1441 GGTCAAATTTATCTCAATATTCTCAGACGCGATACAGGTGAAAAAACCATACAAA 1500
Db 1712 GGTCAAATTTATCTCAATATTCTCAGACGCGATACAGGTGAAAAAACCATACAAA 1771
|||||
QY 1501 TGTGAATTTTGTGAATGCTGCGAGCCAGAGAGCATCTCTGAGGTATCCTTGGAGAGA 1560
Db 1772 TGTGAATTTTGTGAATGCTGCGAGCCAGAGAGCATCTCTGAGGTATCCTTGGAGAGA 1831
|||||

Qy 1561 CATCACAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1620
Db 1832 CATCACAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1891
Qy 1621 GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAACAAATAATTTGAAAGATTT 1680
Db 1892 GACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAACCAATAATTTGAAAGATTT 1951
Qy 1681 TTTGATGGTGCACAAAGATGTTACAGGAGTCCACTGCAAGAGCAGCTTAAAGAGATGCT 1740
Db 1952 TTTGATGGTGCACAAAGATGTTACAGGAGTCCACTGCAAGAGCAGCTTAAAGAGATGCT 2011
Qy 1741 TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCTCTCACCAGACACAAAGATACTCAG 1800
Db 2012 TCTGTTTTTCAAGATGTTCTGGGAGCGCTGTCTCTCACCAGACACAAAGATACTCAG 2071
Qy 1801 GATTTCATAAAATGCAAGCTGATGACAGTGTCTGATAAAGTGAATAAAACCCCTACCCCT 1860
Db 2072 GATTTCATAAAATGCAAGCTGATGACAGTGTCTGATAAAGTGAATAAAACCCCTACCCCT 2131
Qy 1861 GCTTACCTGACCTGTTTAAAAAGAGATCAGAGTTGAAACTCAGGCAATAAATCCTCATC 1920
Db 2132 GCTTACCTGACCTGTTTAAAAAGAGATCAGAGTTGAAACTCAGGCAATAAATCCTCATC 2191
Qy 1921 TGTAGAACCAAGGGGATGTTACTCCTCTCGGATGGCAGTACCACCCATAAATCCTGAA 1980
Db 2192 TGTAGAACCAAGGGGATGTTACTCCTCTCGGATGGCAGTACCACCCATAAATCCTGAA 2251
Qy 1981 GTTAGCCCAAGAGAGAACCAAGGAGACCGAGCTGACTGCAGATACAGGCCAAGTGTG 2040
Db 2252 GTTAGCCCAAGAGAGAACCAAGGAGACCGAGCTGACTGCAGATACAGGCCAAGTGTG 2311
Qy 2041 GATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGCTCTTCAAAATGCGCGCAATTT 2100
Db 2312 GATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGCTCTTCAAAATGCGCGCAATTT 2371
Qy 2101 TCTTTGAGTAAAGTTTCAATTTCAAGTATACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2372 TCTTTGAGTAAAGTTTCAATTTCAAGTATACCTGTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCAGAAAGTTTAAATGATGACACAGAGCTGGAGCATAAATAAATCCTGACGTTTCAAT 2220
Db 2432 TATCAGAAAGTTTAAATGATGACACAGAGCTGGAGCATAAATAAATCCTGACGTTTCAAT 2491
Qy 2221 AAAAATCTGAAACAAAGTCTTGTAGAGTGCAGTACGGATGCCGCCAGCGTTG 2280
Db 2492 AAAAATCTGAAACAAAGTCTTGTAGAGTGCAGTACGGATGCCGCCAGCGTTG 2551
Qy 2281 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAACCCCAAGCCCAAGTCTGCTTTC 2340
Db 2552 CTGGGAAAGATGTCCTCCCTCTCTAGTTTCTGTAACCCCAAGCCCAAGTCTGCTTTC 2611
Qy 2341 CCGGCGCAGTCCAAATCCCTGCTCATCTGCAAGGGGAAAGCAGAGCCCTCTCTGGGCGAGC 2400
Db 2612 CCGGCGCAGTCCAAATCCCTGCTCATCTGCAAGGGGAAAGCAGAGCCCTCTCTGGGCGAGC 2671
Qy 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAAGTCC 2460
Db 2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCAAGTAACTGAAAGTCC 2731
Qy 2461 CACAGACACACAGCAAGTGGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTCAG 2520
Db 2732 CACAGACACACAGCAAGTGGGGGTCCAAAGGGCGCCACAGGCAACAGCAATCTCAG 2791
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAAGACAAAAAGACCCGAGACAAA 2580
Db 2792 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAAGACAAAAAGACCCGAGACAAA 2851
Qy 2581 TTGAAACCTCTTCCAGTAGTCTCTCTCAGGCCACCTCTGGGAGCAGTAAATCAATGGT 2640
Db 2852 TTGAAACCTCTTCCAGTAGTCTCTCTCAGGCCACCTCTGGGAGCAGTAAATCAATGGT 2911

Qy 2641 TCCATCGACTACCCCGCCAAAGAACGACAGCCGCTGGGCACCTCCGGGAAGAGACTATTTTC 2700
Db 2912 TCCATCGACTACCCCGCCAAAGAACGACAGCCGCTGGGCACCTCCGGGAAGAGACTATTTTC 2971
Qy 2701 TGTAAATCGAGTGCAGCAATACTGACAGCAAAATTTGGTGAAGCCCTTCCAAAAGACTG 2760
Db 2972 TGTAAATCGAGTGCAGCAATACTGACAGCAAAATTTGGTGAAGCCCTTCCAAAAGACTG 3031
Qy 2761 AAGTCCAGCGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 2820
Db 3032 AAGTCCAGCGTGGTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTT 3091
Qy 2821 GACCTTCCCAAGTACCATATGTTTGGTGCAGAGCATCATCATCTGTTTACCGCAGGACTGTG 2880
Db 3092 GACCTTCCCAAGTACCATATGTTTGGTGCAGAGCATCATCATCTGTTTACCGCAGGACTGTG 3151
Qy 2881 TATCCGTGCAGGCGCTGCTCCCAAAACCAAGGTTTCTGAGCTCCAGGAGTCTGATTTCT 2940
Db 3152 TATCCGTGCAGGCGCTGCTCCCAAAACCAAGGTTTCTGAGCTCCAGGAGTCTGATTTCT 3211
Qy 2941 CCAAAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGTGGCTCCGGGCCACTTTTACCTTGTG 3000
Db 3212 CCAAAATGTGCTGACTGTTTCAAGAGCCCTATGTTGGTGGCTCCGGGCCACTTTTACCTTGTG 3271
Qy 3001 CCTCTGTGTAGTCCAGCATCCAGCTCCAGCTTACGCTGAGTCCAGGAGTCTGATTTCT 3034
Db 3272 CCTCTGTGTAGTCCAGCATCCAGCTCCAGCTTACGAGTCCAGGAGGAAAGGCTGTGTCATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGTGGCACAAGAGAACTATGAGAAATTTTATTTGGGAATGCAAT 3391
Qy 3035 ----- 3034
Db 3392 TATCGACCAATGACAAAAAACTTGATTTCACTAAATAGGGGAAAAAAGGCTTTGGTGG 3451
Qy 3048 ATGTGAGTGTCTTACTCCCATGAAATTTTACTTTTACTTTTCTTGAAGAGCGAATGGT 3107
Db 3452 ATGTGAGTGTCTTACTCCCATGAAATTTTACTTTTACTTTTCTTGAAGAGCGAATGGT 3511
Qy 3108 GAAAGCTACTGAAATAGTGTGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 3167
Db 3512 GAAAGCTACTGAAATAGTGTGATTTGATGTTGATGTTGATGTTGATGTTGATGTTGATGTTG 3571
Qy 3168 GAACACTACAGTCTGTGTA 3186
Db 3572 GAACACTACAGTCTGTGTA 3590

RESULT 4

US-09-949-016-4262
; Sequence 4262, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4262
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4262

Query Match	95.4%	Score 3040.6	DB 4	Length 5632
Best Local Similarity	95.8%	Pred. No. 0		
Matches 3180	Conservative 6	Mismatches 0	Indels 133	Gaps 1
Qy	1	ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG	60	
Db	272	ATGCNATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAATGTACATGGATGGG	331	
Qy	61	CCGAAAGTGATGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG	120	
Db	332	CCGAAAGTGATGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGGATGCTTGTCAATG	391	
Qy	121	AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCACAAATCGAG	180	
Db	392	AAAGGGACCGCTGTGTTCATTCGAGCTACACAAGAAAAAATGTCATCCAAATCGAG	451	
Qy	181	GGGTATATGCCCTTGGATTCGATGTTCTCGAGCCAGACCTTCACACATTCAGAGACCTT	240	
Db	452	GGGTATATGCCCTTGGATGTCATGTTCTCGAGCCAGACCTTCACACATTCAGAGACCTT	511	
Qy	241	AATAAAACATGTCTTAATGCAACCGCGCTACCCCTCTGTGAACCAGCAGTTCCTCGGGTT	300	
Db	512	AATAAAACATGTCTTAATGCAACCGCGCTACCCCTCTGTGAACCAGCAGTTCCTCGGGTT	571	
Qy	301	GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACCTCCCAAGGAA	360	
Db	572	GAAGCAGAGTATCTCAGTCCGCTTGATAAAGTCAAGTGCAGAACCTCCCAAGGAA	631	
Qy	361	AAGNATTCGAAGNAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGATCGCT	420	
Db	632	AAGNATTCGAAGGAAAAATGAATTTAGCTGTGAGTATGTGGCAGACATTTAGATCGCT	691	
Qy	421	TTTGATGTGAGATGCACATGAGAAACACAAAGATTCCTTCACTACGGGTGTAAATG	480	
Db	692	TTTGATGTGAGATGCACATGAGAAACACAAAGATTCCTTCACTACGGGTGTAAATG	751	
Qy	481	TGCGGAAGAAGMTTSRRSSAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGC	540	
Db	752	TGCGGAAGAAGMTTCAAGGAGCCTTGGTTTCTTAAAAATCACATGCGGACACATAATGGC	811	
Qy	541	AAATCGGGGGCCAGAACCAACTGCACAGAGCTTTGAGAGTAGTTCAGCAACGATCAAC	600	
Db	812	AAATCGGGGGCCAGAACCAACTGCACAGAGCTTTGAGAGTAGTTCAGCAACGATCAAC	871	
Qy	601	GAGTCTGTCCAGGTGCACGGCGCGAGAGCATCTCCTCTCTTACAAAATCTGCATGTT	660	
Db	872	GAGTCTGTCCAGGTGCACGGCGCGAGAGCATCTCCTCTCTTACAAAATCTGCATGTT	931	
Qy	661	TGTGGCTTCCTATTTCAAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA	720	
Db	932	TGTGGCTTCCTATTTCAAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACAAA	991	
Qy	721	AAAACCTGTTTCGGTACAGCAGCGCGAGACAGACTCTCCACAAGGAGGAATGCCGTCC	780	
Db	992	AAAACCTGTTTCGGTACAGCAGCGCGAGACAGACTCTCCACAAGGAGGAATGCCGTCC	1051	
Qy	781	TCGAGGAGGAGCTTCCTGCAAGTGTTCAACTTGAGACCAAACTCTACCTGTAACCGGG	840	
Db	1052	TCGAGGAGGAGCTTCCTGCAAGTGTTCAACTTGAGACCAAACTCTACCTGTAACCGGG	1111	
Qy	841	AAGAAGCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG	900	
Db	1112	AAGAAGCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCACCTTCCAGGCTTGGCAG	1171	
Qy	901	CTGCTTACAAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG	960	
Db	1172	CTGCTTACAAAGGAAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGGCAAGAGGG	1231	
Qy	961	AGCACCGACAAACGACGATTCGAGTTCGAGNAGAGCTTGGAGAAAACAAATAGGGCAGT	1020	
Db	1232	AGCACCGACAAACGACGATTCGAGTTCGAGNAGAGCTTGGAGAAAACAAATAGGGCAGT	1291	

QY	1021	TGTCAGGCGCTCTCGCAAGAGAAAGAGAGTGC	1081	TTGTCAGGCGCTCTCGCAAGAGAAAGAGAGTGC
DB	1292	TGTCAGGCGCTCTCGCAAGAGAAAGAGAGTGC	1351	TGTCAGGCGCTCTCGCAAGAGAAAGAGAGTGC
QY	1081	GTGGACCGCGATCCCAAGTTCACCGATGAC	1140	GTGGACCGCGATCCCAAGTTCACCGATGAC
DB	1352	GTGGACCGCGATCCCAAGTTCACCGATGAC	1411	GTGGACCGCGATCCCAAGTTCACCGATGAC
QY	1141	GGCAAAAGCTTTTCAGAACTTACCAACAGCT	1200	GGCAAAAGCTTTTCAGAACTTACCAACAGCT
DB	1412	GGCAAAAGCTTTTCAGAACTTACCAACAGCT	1471	GGCAAAAGCTTTTCAGAACTTACCAACAGCT
QY	1201	CGGAGGCGCGCGCGAGTGC	1260	CGGAGGCGCGCGCGAGTGC
DB	1472	CGGAGGCGCGCGCGAGTGC	1531	CGGAGGCGCGCGCGAGTGC
QY	1261	TCTCCTGACCTCGCGCGCGCTCTGATG	1320	TCTCCTGACCTCGCGCGCGCTCTGATG
DB	1532	TCTCCTGACCTCGCGCGCGCTCTGATG	1591	TCTCCTGACCTCGCGCGCGCTCTGATG
QY	1321	TCTGAAGACCGGATCTGAGGATGGCTTCC	1380	TCTGAAGACCGGATCTGAGGATGGCTTCC
DB	1592	TCTGAAGACCGGATCTGAGGATGGCTTCC	1651	TCTGAAGACCGGATCTGAGGATGGCTTCC
QY	1381	GGAGGAAAAATAAACAATCTTCAACATCTT	1440	GGAGGAAAAATAAACAATCTTCAACATCTT
DB	1652	GGAGGAAAAATAAACAATCTTCAACATCTT	1711	GGAGGAAAAATAAACAATCTTCAACATCTT
QY	1441	CGTTCAAAATTTTACCTCAATATTTCACTC	1500	CGTTCAAAATTTTACCTCAATATTTCACTC
DB	1712	CGTTCAAAATTTTACCTCAATATTTCACTC	1771	CGTTCAAAATTTTACCTCAATATTTCACTC
QY	1501	TGTTGAAATTTTGTGAATATGCTGCGAGCC	1560	TGTTGAAATTTTGTGAATATGCTGCGAGCC
DB	1772	TGTTGAAATTTTGTGAATATGCTGCGAGCC	1831	TGTTGAAATTTTGTGAATATGCTGCGAGCC
QY	1561	CATCAAGGAAAAAACAACCGATGTTGCTG	1620	CATCAAGGAAAAAACAACCGATGTTGCTG
DB	1832	CATCAAGGAAAAAACAACCGATGTTGCTG	1891	CATCAAGGAAAAAACAACCGATGTTGCTG
QY	1621	GACACTGAAGTGCACATTTAAACCGCTG	1680	GACACTGAAGTGCACATTTAAACCGCTG
DB	1892	GACACTGAAGTGCACATTTAAACCGCTG	1951	GACACTGAAGTGCACATTTAAACCGCTG
QY	1681	TTTTCATGCTGCAAGAGTGTTCAGGAGTCC	1740	TTTTCATGCTGCAAGAGTGTTCAGGAGTCC
DB	1952	TTTTCATGCTGCAAGAGTGTTCAGGAGTCC	2011	TTTTCATGCTGCAAGAGTGTTCAGGAGTCC
QY	1741	TCTGTTTTTCAGAAATGTTTCTGGGAGCG	1800	TCTGTTTTTCAGAAATGTTTCTGGGAGCG
DB	2012	TCTGTTTTTCAGAAATGTTTCTGGGAGCG	2071	TCTGTTTTTCAGAAATGTTTCTGGGAGCG
QY	1801	GATTTTCATAAATAATGACGTGTGATTAAG	1860	GATTTTCATAAATAATGACGTGTGATTAAG
DB	2072	GATTTTCATAAATAATGACGTGTGATTAAG	2131	GATTTTCATAAATAATGACGTGTGATTAAG
QY	1861	GCTTTACCTGACCTGTTTAAAAAGAGATC	1920	GCTTTACCTGACCTGTTTAAAAAGAGATC
DB	2132	GCTTTACCTGACCTGTTTAAAAAGAGATC	2191	GCTTTACCTGACCTGTTTAAAAAGAGATC
QY	1921	TGTTAGAACCAAGCGGATGTTTACTCTCTC	1980	TGTTAGAACCAAGCGGATGTTTACTCTCTC
DB	2192	TGTTAGAACCAAGCGGATGTTTACTCTCTC	2251	TGTTAGAACCAAGCGGATGTTTACTCTCTC
QY	1981	GTTTAGCCCCAAGAGAGCAAAACCGAGAC	2040	GTTTAGCCCCAAGAGAGCAAAACCGAGAC
DB	2252	GTTTAGCCCCAAGAGAGCAAAACCGAGAC	2311	GTTTAGCCCCAAGAGAGCAAAACCGAGAC
QY	2041	GATTTGTCAAGAAAAACCTTTTAAATTTAT	2100	GATTTGTCAAGAAAAACCTTTTAAATTTAT
DB	2312	GATTTGTCAAGAAAAACCTTTTAAATTTAT	2371	GATTTGTCAAGAAAAACCTTTTAAATTTAT
QY	2101	TCTTTTGTAGTAAAAAGTTTTGATTTCCA	2160	TCTTTTGTAGTAAAAAGTTTTGATTTCCA

Db 2372 TCTTTGAGTAAAGTTTGTATTTCCAAAGTATCACCTGTCTCCATTTTGTACCTTCAAGACATTT 2431
Qy 2161 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATCAATCTCGAGTTTCAT 2220
Db 2432 TATCCAGAGTTTAAATGATGACACAGAGACTGGAGCATAAATCAATCTCGAGTTTCAT 2491
Qy 2221 AAAAATCTCGAAACCAAGTCTTGTGTAGAGTTCGAGTACCGGATGCCGCCAGCGTTG 2280
Db 2492 AAAAATCTCGAAACCAAGTCTTGTGTAGAGTTCGAGTACCGGATGCCGCCAGCGTTG 2551
Qy 2281 CTGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2340
Db 2552 CTGGAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACCAAGCCCAAGTCTGCTTTTC 2611
Qy 2341 CCGGCGAGTCCAAATCCCTGCCATCTCGAAGGGGAGCAGAGCCCTCTGGGCCAGGC 2400
Db 2612 CCGGCGAGTCCAAATCCCTGCCATCTCGAAGGGGAGCAGAGCCCTCTGGGCCAGGC 2671
Qy 2401 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2460
Db 2672 AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAGTCC 2731
Qy 2461 CACAGACACACAGAGAATGTGGGGTCAAGGGGCCGCCACACAGCCCAACCAATCTGAG 2520
Db 2732 CACAGACACACAGAGAATGTGGGGTCAAGGGGCCGCCACACAGCCCAACCAATCTGAG 2791
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAAGACCCGAGACAAA 2580
Db 2792 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGATAGACAAAAGACCCGAGACAAA 2851
Qy 2581 TTGAAACCTCTTCAGTAGTCTCTTCTCAGGCCACCTCGGACAGTAAATCAATGGT 2640
Db 2852 TTGAAACCTCTTCAGTAGTCTCTTCTCAGGCCACCTCGGACAGTAAATCAATGGT 2911
Qy 2641 TCCATCGACTACCCGCCAAGAACGACAGCCCGTGGGCACCTCCGGGAAGAGACTATTTC 2700
Db 2912 TCCATCGACTACCCGCCAAGAACGACAGCCCGTGGGCACCTCCGGGAAGAGACTATTTC 2971
Qy 2701 TGTAACTCGAGTCCAGCAATACTGACAGAGAAATTTGTGTAGCCCTTCCAAAAGACTG 2760
Db 2972 TGTAACTCGAGTCCAGCAATACTGACAGAGAAATTTGTGTAGCCCTTCCAAAAGACTG 3031
Qy 2761 AAGTCCAGCTGTGTTGCCCTTGACCTTCAGCCAGCCCGGGCCCAATTCAGAAAGAGCTAT 2820
Db 3032 AAGTCCAGCTGTGTTGCCCTTGACCTTCAGCCAGCCCGGGCCCAATTCAGAAAGAGCTAT 3091
Qy 2821 GACCTTCCCAAGTACCATATGTTGACAGGCAATCAATCACTGTTTACCAGGACTGTGTG 2880
Db 3092 GACCTTCCCAAGTACCATATGTTGACAGGCAATCAATCACTGTTTACCAGGACTGTGTG 3151
Qy 2881 TATCCGTCCAGGCGCTGCTCCAAACCAAGGTTCTGAGCTCCAGCGAGTGCATTTCT 2940
Db 3152 TATCCGTCCAGGCGCTGCTCCAAACCAAGGTTCTGAGCTCCAGCGAGTGCATTTCT 3211
Qy 2941 CCAATGTGCTGACTGTTCAGAAAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAATGTGCTGACTGTTCAGAAAGCCCTATGTTGGCTCCGGGCCACTTTACACTTGTGTG 3271
Qy 3001 CCTGCTGTGAGTCCAGCAATCCAGCTCGACGTTAG----- 3034
Db 3272 CCTGCTGTGAGTCCAGCAATCCAGCTCGACGTTAGAGAGAAAGGCGCTGTGTGCATATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGGCAAAAGAGAAACTATAGAAATTTTATTTGGGAATGCAT 3391
Qy 3035 ----- 3034
Db 3392 TATCGACCAATGACAAAAAATTTGATTCTACTAATTTAGGGGAAAAAGGCTTTGGTGG 3451
Qy 3048 ATGTCAGTGTCTTATCCCATGAAATTAATTTTACTTCACTCTTTGAGAGCGAATGGT 3107

Db 3452 ATGTCAGTGTCTTACTCCCATGAAATTAATTTTACTTCTCATCTTTGAGAGCGAATGGT 3511
Qy 3108 GAAAGCTACTGAAATAAGCTGTGATTTCTACTGTACATAAAACATATCAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATAAGCTGTGATTTCTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
Qy 3168 GAACACTACAGTTGTGTAA 3186
Db 3572 GAACACTACAGTTGTGTAA 3590
RESULT 5
US-09-949-016-12604
; Sequence 12604, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12604
; LENGTH: 20022
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12604

Query Match 49.0%; Score 1561.2; DB 4; Length 20022;
Best Local Similarity 99.5%; Pred. No. 0;

Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 1482 AGGTGAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 1541
Db 7816 AGGTGAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 7875
Qy 1542 GAGGTATCATTGAGAGACATCAAGAGAAAAACCAACCATGTTGCTGCTGAAGTCAA 1601
Db 7876 GAGGTATCATTGAGAGACATCAAGAGAAAAACCAACCATGTTGCTGCTGAAGTCAA 7935
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCCTATTAAACCGCTGACAGTGCAGAAC 1661
Db 7936 GAACGATGGTAAAAATCAGGACACTGAAGATGCCTATTAAACCGCTGACAGTGCAGAAC 7995
Qy 1662 CAAAAATTTGAAAGATTTTGTGATGTTGTCAGAAAGTCTTACAGGAGTCCACCTGCATA 1721
Db 7996 CAAAAATTTGAAAGATTTTGTGATGTTGTCAGAAAGTCTTACAGGAGTCCACCTGCATA 8055
Qy 1722 GCAGCTTAAGAGATGCTTCTGTTTTCAGAAAGTCTTGGGCGAGCGTGTCTCTCACC 1781
Db 8056 GCAGCTTAAGAGATGCTTCTGTTTTCAGAAAGTCTTGGGCGAGCGTGTCTCTCACC 8115
Qy 1782 AGCACAAAGATCTCAGGATTTCCATAAAATGACAGCTGATGACAGTGTGATAAAGT 1841
Db 8116 AGCACAAAGATCTCAGGATTTCCATAAAATGACAGCTGATGACAGTGTGATAAAGT 8175
Qy 1842 GAATAAAACCTACCTCCCTGCTTACCTGAGCTGTAAAAAGATCAGCAGTTGAAAC 1901
Db 8176 GAATAAAACCTACCTCCCTGCTTACCTGAGCTGTAAAAAGATCAGCAGTTGAAAC 8235
Qy 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 1961
Db 8236 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAG 8295
Qy 1962 TACCACCATTAACCTTGAAGTTAGCCCCAAGAGAACAAACGGAGACCGCAGCTGACTG 2021

Db 8296 TACCACCCATACCTTGAAGTTAGCCCAAGAGAAAGCAACCGAGACCGAGCTGACTG 8355
Qy 2022 CAGATACAGGCCAAGTGTGGATTGTCAAGAAAACTTTAAATTTATCGTGGGGGCTCT 2081
Db 8356 CAGATACAGGCCAAGTGTGGATTGTCAAGAAAACTTTAAATTTATCGTGGGGGCTCT 8415
Qy 2082 TCACAAATTCGCCGCAATTTCTTCAGTAAGATTTGATTCAGATATCACTGTCCTATT 2141
Db 8416 TCACAAATTCGCCGCAATTTCTTCAGTAAGATTTGATTCAGATATCACTGTCCTATT 8475
Qy 2142 TTGTACCTTCAAGACATTTATCCAGAAGTTTAAATGATGCACACAGAGCTGGACATAA 2201
Db 8476 TTGTACCTTCAAGACATTTATCCAGAAGTTTAAATGATGCACACAGAGCTGGACATAA 8535
Qy 2202 ATACATCTGACGTTTCATAAAAACTGTGGAACCAAGTCTTGTGTTAGAAATCGAGTAC 2261
Db 8536 ATACATCTGACGTTTCATAAAAACTGTGGAACCAAGTCTTGTGTTAGAAATCGAGTAC 8595
Qy 2262 CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTCTTCTGTAACCC 2321
Db 8596 CGGATGCCCGCAGCGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTCTTCTGTAACCC 8655
Qy 2322 CAAGCCCAAGTGTGCTTTCCCGCGCAGTCCAAATCCCTGCTGCAATCTGCAAGGGGAAGCA 2381
Db 8656 CAAGCCCAAGTGTGCTTTCCCGCGCAGTCCAAATCCCTGCTGCAATCTGCAAGGGGAAGCA 8715
Qy 2382 GAGCCCTCTGCGGCCAGGCAAGCCCTCTGATCTCAGGGATGAGACTTAGCACTTTAGC 2441
Db 8716 GAGCCCTCTGCGGCCAGGCAAGCCCTCTGATCTCAGGGATGAGACTTAGCACTTTAGC 8775
Qy 2442 CCCAGTAACTGAAGTCCACAGACCAAGAGATGTGGGGTCCCAAGGGGCGCCAC 2501
Db 8776 CCCAGTAACTGAAGTCCACAGACCAAGAGATGTGGGGTCCCAAGGGGCGCCAC 8835
Qy 2502 CAGGCAACAGCAATCTGAGATGTTTCTAAACCAAGTGTTCCTCCCTGCAACCGGATAAGAC 2561
Db 8836 CAGGCAACAGCAATCTGAGATGTTTCTAAACCAAGTGTTCCTCCCTGCAACCGGATAAGAC 8895
Qy 2562 AAAAGACCCGAGACAAAATTTGAAACCTTTCCAGTGTCTCTTCCAGCCCACTCCG 2621
Db 8896 AAAAGACCCGAGACAAAATTTGAAACCTTTCCAGTGTCTCTTCCAGCCCACTCCG 8955
Qy 2622 CAGCAGTAACTAATGTTCCATGACTACCCCGCAAGACGACGACCCGCTGGGAC 2681
Db 8956 CAGCAGTAACTAATGTTCCATGACTACCCCGCAAGACGACGACCCGCTGGGAC 9015
Qy 2682 TCCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGAGAAATTTGGTGA 2741
Db 9016 TCCGGGAAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGAGAAATTTGGTGA 9075
Qy 2742 GCCCTTCCAAAAGACTGAAGTCCAGCGTGTGCTGCTTCCAGTGTGACGCTTCCAGCCCGGGC 2801
Db 9076 GCCCTTCCAAAAGACTGAAGTCCAGCGTGTGCTGCTTCCAGTGTGACGCTTCCAGCCCGGGC 9135
Qy 2802 CAATTAAGAGAGAGCTATGACTTCCCAAGTACCATATGTTGAGAGATCAATCACT 2861
Db 9136 CAATTAAGAGAGAGCTATGACTTCCCAAGTACCATATGTTGAGAGATCAATCACT 9195
Qy 2862 GTTACCGCAGGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGTTCCTGAG 2921
Db 9196 GTTACCGCAGGACTGTGTATCCGTGCGAGCGCTGCTCCCAACCAAGTTCCTGAG 9255
Qy 2922 CTCAGCGAGGTGCAATTTCCAAATGTGCTGATGTTTCCAGAGCCCTATGTTGGCTCCGG 2981
Db 9256 CTCAGCGAGGTGCAATTTCCAAATGTGCTGATGTTTCCAGAGCCCTATGTTGGCTCCGG 9315
Qy 2982 GCCACTTTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
Db 9316 GCCACTTTACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9375
Qy 3042 TGGTGGATGTCAGT 3055

Db 9376 TGCATGAGGGGGT 9389
RESULT 6
US-09-949-016-16004
; Sequence 16004, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16004
; LENGTH: 20023
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-16004
Query Match 49.0%; Score 1561.2; DB 4; Length 20023;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTGTGAATATGCTGCAGCCCAAGACATCTCT 1541
Db 7816 AGGTGAAAAACCATACAAATGTGAATTTGTGAATATGCTGCAGCCCAAGACATCTCT 7875
Qy 1542 GAGTATCATCTTGAGAGACATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAA 1601
Db 7876 GAGTATCATCTTGAGAGACATCAAGGAAAAAACAACCGATGTTGCTGCTGAAGTCAA 7935
Qy 1602 GAACGATGTTAAAAATCAGGACACTGAAGATGCTACTATTAAACCGCTGACGTCGCAAC 1661
Db 7936 GAACGATGTTAAAAATCAGGACACTGAAGATGCTACTATTAAACCGCTGACGTCGCAAC 7995
Qy 1662 CAAAAATTTGAAAAAGATTTTGTGATGTTGTCGCAAGATGTTTACAGGAGTCCACCTGCAAA 1721
Db 7996 CAAAAATTTGAAAAAGATTTTGTGATGTTGTCGCAAGATGTTTACAGGAGTCCACCTGCAAA 8055
Qy 1722 GCAGCTTAAGAGATGCTCTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 1781
Db 8056 GCAGCTTAAGAGATGCTCTCTGTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 8115
Qy 1782 AGCACACAAAGATACTCAGGATTTTCCATAAAATGACAGTGTGACAGTGTGATAAAGT 1841
Db 8116 AGCACACAAAGATACTCAGGATTTTCCATAAAATGACAGTGTGACAGTGTGATAAAGT 8175
Qy 1842 GAATAAAAACCTTACCCTGCTTACCTGGACCTCTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8176 GAATAAAAACCTTACCCTGCTTACCTGGACCTCTTAAAAAAGAGATCAGCAGTTGAAAC 8235
Qy 1902 TCAGGCAAAATACCTCATCTGTAGAACCAAGGCGGATGTTTACTCTCTCCGGATGGAG 1961
Db 8236 TCAGGCAAAATACCTCATCTGTAGAACCAAGGCGGATGTTTACTCTCTCCGGATGGAG 8295
Qy 1962 TACCACCCATACCTTGAAGTTAGCCCAAGAGACAAACCGAGACCGGAGCTGACTG 2021
Db 8296 TACCACCCATACCTTGAAGTTAGCCCAAGAGACAAACCGAGACCGGAGCTGACTG 8355
Qy 2022 CAGATACAGGCCAAGTGTGATTTGTCACGAAAAACCTTTAAATTTATCCGCTGGGGCTCT 2081
Db 8356 CAGATACAGGCCAAGTGTGATTTGTCACGAAAAACCTTTAAATTTATCCGCTGGGGCTCT 8415
Qy 2082 TCACAAATTCGCCGCAATTTCTTTGAGTAAAGATTTGATTCACCAAGTATCACCTGTCATT 2141

Db 8416 TCACAAATGCCCGCAATTTCTTTGAGTAAAGTTTGATTCACAGTATCACTGTCCATT 8475
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGACTGGAGCATAA 2201
Db 8476 TTGTACCTTCAAGACATTTTATCCAGAGTTTATATGATGACACAGAGACTGGAGCATAA 8535
Qy 2202 ATACAATCCTGACGTTTCATAAAAACTGTCGAAAACAAAGTCTTGTGTTAGAACTGACAGTAC 2261
Db 8536 ATACAATCCTGACGTTTCATAAAAACTGTCGAAAACAAAGTCTTGTGTTAGAACTGACAGTAC 8595
Qy 2262 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 2321
Db 8596 CGGATGCCCGCCAGCGTTGCTGGGAAAAGATGTGCTCCCTCTCTAGTTTCTGTAAACC 8655
Qy 2322 CAAGCCCAAGTCTGTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 2381
Db 8656 CAAGCCCAAGTCTGTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCA 8715
Qy 2382 GAGCCCTCCTGGGCCAGGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 2441
Db 8716 GAGCCCTCCTGGGCCAGGCAAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGC 8775
Qy 2442 CCCAAGTTAACTGAAGTCCCAAGACACAGAGAAATGTGGGGTCCAAAGGGGCGCCAC 2501
Db 8776 CCCAAGTTAACTGAAGTCCCAAGACACAGAGAAATGTGGGGTCCAAAGGGGCGCCAC 8835
Qy 2502 CAGCMAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGCCACCGGATAGAC 2561
Db 8836 CAGGCAACAGCAATCTGAGATGTTTCTTAAACAGAGTGTTCCTCCCTGCCACCGGATAGAC 8895
Qy 2562 AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTCTCAGCCCAACCTCGG 2621
Db 8896 AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCTCTCAGCCCAACCTCGG 8955
Qy 2622 CAGCAGTAACATCAATGTTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 2681
Db 8956 CAGCAGTAACATCAATGTTTCATCGACTACCCCGCCAAAGAACGACAGCCCGTGGGCACC 9015
Qy 2682 TCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGCAATCTGCGAGCAATTTGGTGA 2741
Db 9016 TCCGGGAGAGACTATTTCTGTAATCGAGTGCAGCAATCTGCGAGCAATTTGGTGA 9075
Qy 2742 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTTGACCGCCCGGGGC 2801
Db 9076 GCCCCTTCCAAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTTGACCGCCCGGGGC 9135
Qy 2802 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCAATATGTTGTCAGAGGCATCAATCACT 2861
Db 9136 CAATTACAGAGAGGCTATGACCTTCCCAAGTACCAATATGTTGTCAGAGGCATCAATCACT 9195
Qy 2862 GTTACCGCAGGACTGTGTATCCGTCCAGCGCGCTGCCCTCCCAACCAAGGTTCCCTGAG 2921
Db 9196 GTTACCGCAGGACTGTGTATCCGTCCAGCGCGCTGCCCTCCCAACCAAGGTTCCCTGAG 9255
Qy 2922 CTCAGCGAGGTCGATTCTCAAAATGTGCTGACTGTTCAGAGCCCTATGTTGGTCCGG 2981
Db 9256 CTCAGCGAGGTCGATTCTCAAAATGTGCTGACTGTTCAGAGCCCTATGTTGGTCCGG 9315
Qy 2982 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTTAGAAAGTCT 3041
Db 9316 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTTAGAAAGTAT 9375
Qy 3042 TGGTGGATGTCAGT 3055
Db 9376 TGCATGAGGGGCGT 9389

RESULT 7
US-08-892-695-9
; Sequence 9, Application US/0892695A
; Patent No. 6808878
; GENERAL INFORMATION:

; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowal, David
; APPLICANT: Kommen, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892.695A
; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 10365
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Genomic
; OTHER INFORMATION: Sequence encoding ZABC1
; Patent No. 6808878
US-08-892-695-9

Query Match 49.0%; Score 1559.6; DB 4; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGTCTCAGCCACAGAGACATCTCT 1541
Db 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGTCTCAGCCACAGAGACATCTCT 8345
Qy 1542 GAGTATCACTTGTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGTGAAGTCAA 1601
Db 8346 GAGTATCACTTGTGAGAGACATCAAGAGAAAAACAAACCGATGTTGCTGTGAAGTCAA 8405
Qy 1602 GAACGATGTTAAATCAGGACACTGAAGTGCACCTATTAAACCGCTGACAGTGCACAAAC 1661
Db 8406 GAACGATGTTAAATCAGGACACTGAAGTGCACCTATTAAACCGCTGACAGTGCACAAAC 8465
Qy 1662 CAAAAATTTGAAAAAGATTTTTTGTGTTGTCGCAAAAGATGTTACAGGACGTCACCTGCAAA 1721
Db 8466 CAAAAATTTGAAAAAGATTTTTTGTGTTGTCGCAAAAGATGTTACAGGACGTCACCTGCAAA 8525
Qy 1722 GCAGCTTAAGAGATGCCCTCTGTTTTTCAGAAATGTTCTGGGCGAGCGCTCTCTCACC 1781
Db 8526 GCAGCTTAAGAGATGCCCTCTGTTTTTCAGAAATGTTCTGGGCGAGCGCTCTCTCACC 8585
Qy 1782 AGCACACAAAGATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
Db 8586 AGCACACAAAGATCTCAGGATTTCCATAAAATGCAGCTGATGACAGTGTGATAAAGT 8645
Qy 1842 GAATAAAAAACCTTACCCTGCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8646 GAATAAAAAACCTTACCCTGCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 8705
Qy 1902 TCAGGCAATTAACCTCATCTGTAGAAACCAAGGCGGATGTTACTCTCTCCGATGGGAG 1961
Db 8706 TCAGGCAATTAACCTCATCTGTAGAAACCAAGGCGGATGTTACTCTCTCCGATGGGAG 8765
Qy 1962 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGCAACCGAGACCGGAGCTGACTG 2021
Db 8766 TACCACCCATAACCTTGAAGTTAGCCCAAGAGAGCAACCGAGAGCTGACTG 8825
Qy 2022 CAGATACAGGCCAAAGTGTGGATTGTTCACGAAAAACCTTTTAAATTTTATCCGTGGGGCTCT 2081
Db 8826 CAGATACAGGCCAAAGTGTGGATTGTTCACGAAAAACCTTTTAAATTTTATCCGTGGGGCTCT 8885
Qy 2082 TCACAAATGGCCGGCAATTTCTTTGAGTAAAGTTTGAATTCAGAGTATCACTGTCCATT 2141

Db	8886	TCACAAATTGCCCGCAATTTCTTTGAGTAAAGTTTGATTCCAAGTATACCTGTGTCATT	8945
Qy	2142	TTGTACTCTTCAAGACAATTTTATCCAGAAGTTTTTAATGATCACACAGAGACTGGAGCATAA	2201
Db	8946	TTGTACTCTTCAAGACAATTTTATCCAGAAGTTTTTAATGATGACCAGAGACTGGAGCATAA	9005
Qy	2202	ATACAAATCCTGACGTTTCATAAAACTCTCGAAACAAGTCCCTTGCCTTAGAAGTCGACGTAC	2261
Db	9006	ATACAAATCCTGACGTTTCATAAAACTCTCGAAACAAGTCCCTTGCCTTAGAAGTCGACGTAC	9065
Qy	2262	CGGATGCCCGCCAGCGTTTGTCTGGGAAAGATGTGCCTCCCTCTCTAGTTTCTGTAAACC	2321
Db	9066	CGGATGCCCGCCAGCGTTTGTCTGGGAAAGATGTGCCTCCCTCTCTAGTTTCTGTAAACC	9125
Qy	2322	CAAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGTCATCTCGGAAGGGGAAGCA	2381
Db	9126	CAAGCCCAAGTCTGCTTTCCCGGCGAGTCCAAATCCCTGTCATCTCGGAAGGGGAAGCA	9185
Qy	2382	GAGCCCTCTTGGGCGAGCAAGGCCCTCTGTGACTTCAGGGATAGACTCTAGCACTTTTACG	2441
Db	9186	GAGCCCTCTTGGGCGAGCAAGGCCCTCTGTGACTTCAGGGATAGACTCTAGCACTTTTACG	9245
Qy	2442	CCCAAGTAACTGAAGTCCACAGACACAGCAGAATGTGGGGTCCAAAGGGCCCGCCAC	2501
Db	9246	CCCAAGTAACTGAAGTCCACAGACACAGCAGAATGTGGGGTCCAAAGGGCCCGCCAC	9305
Qy	2502	CAGGCAACAGCAATCTCAGATGTTTCTTAAACCAGTGTTTCCCTCTGACCGGATAGAC	2561
Db	9306	CAGGCAACAGCAATCTCAGATGTTTCTTAAACCAGTGTTTCCCTCTGACCGGATAGAC	9365
Qy	2562	AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCCTTCTCAGGCCACCTCGG	2621
Db	9366	AAAAAGACCCGAGACAAAATTTGAAACCTCTTCCAGTAGCTCCTTCTCAGGCCACCTCGG	9425
Qy	2622	CAGCAGTAACTCAATGGTTCATCGACTACCCGCCCAAGAACGACGCCGTGGGCACC	2681
Db	9426	CAGCAGTAACTCAATGGTTCATCGACTACCCGCCCAAGAACGACGCCGTGGGCACC	9485
Qy	2682	TCCGGGAAGAGACTAATTTCTGTAATCGGAGTGCCAGCAATACTTGCAGCAGAAATTTGGTGA	2741
Db	9486	TCCGGGAAGAGACTAATTTCTGTAATCGGAGTGCCAGCAATACTTGCAGCAGAAATTTGGTGA	9545
Qy	2742	GCCCTTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGAGCTTGACAGCCCGGGGC	2801
Db	9546	GCCCTTTCCAAAAGACTGAAGTCCAGCGTGGTTGCCCTTGAGCTTGACAGCCCGGGGC	9605
Qy	2802	CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT	2861
Db	9606	CAATTACAGAAAGGCTATGACCTTCCCAAGTACCATATGGTCAGAGGCATCACATCACT	9665
Qy	2862	GTTAACCGAGGACTGTGTATCCGTGCGAGGCGCTGCCCTCCCAAAACCAAGGTTCTGTAG	2921
Db	9666	GTTAACCGAGGACTGTGTATCCGTGCGAGGCGCTGCCCTCCCAAAACCAAGGTTCTGTAG	9725
Qy	2922	CTCCAGCGAGGTGCAATTTCCAAATGTGCTGACTGTTCAGAAAGCCCTATGGTGGCTCCGG	2981
Db	9726	CTCCAGCGAGGTGCAATTTCCAAATGTGCTGACTGTTCAGAAAGCCCTATGGTGGCTCCGG	9785
Qy	2982	GCCACTTTACACTTGTGTGCTGCTGGTAGTCCAGCATCCAGCTCGACGTTTGAAGGTCT	3041
Db	9786	GCCACTTTACACTTGTGTGCTGCTGGTAGTCCAGCATCCAGCTCGACGTTTGAAGGTAT	9845
Qy	3042	TGTTGGATGTCACT 3055	
Db	9846	TGATGAGGGCGT 9859	

APPLICANT: Gray, Joe W.
 APPLICANT: Collins, Colin
 APPLICANT: Hwang, Soo-in
 APPLICANT: Godfrey, Tony
 APPLICANT: Kowbel, David
 APPLICANT: Rommens, Johanna
 TITLE OF INVENTION: Genes from the 20ql3 Amplicon and Their
 TITLE OF INVENTION: Uses
 NUMBER OF SEQUENCES: 40
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Townsend and Townsend and Crew LLP
 STREET: Two Embarcadero Center, Eighth Floor
 STATE: San Francisco
 COUNTRY: USA
 ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/680,395
 FILING DATE: 15-JUL-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Bastian, Kevin L.
 REGISTRATION NUMBER: 34,774
 REFERENCE/DOCKET NUMBER: 023070-06890005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 576-0200
 TELEFAX: (415) 576-0300
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1507 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: -
 LOCATION: 1..1507
 OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
 OTHER INFORMATION: transcript with homology to C2H2 zinc
 OTHER INFORMATION: finger genes"
 US-08-680-395-3

Query Match	34.7%;	Score 1106.2;	DB 2;	Length 1507;
Best Local Similarity	96.6%;	Pred. No. 0;		
Matches 1150;	Conservative 9;	Mismatches 28;	Indels 3;	Gaps 3;
Qy	1	ATGCAATCGAAAGTGACAGGAACAATGCCAACTCAATCCCTCTTAAATGTACATGATGGG	60	
Dd	320	ATGCAATCGAAAGTGACAGGAACAATGCCAACTCAATCCCTCTTAAATGTACATGAT-GG	378	
Qy	61	CCAGAACTGATTGCCAGCTCTCTTTGGCAGTCGAGATGGAGATGGCGCTTGTTCGAATG	120	
Dd	379	CCAAGACTGTAATGGCAGCTCTCTTTGCCAGTCGAGATGGAGATGGA-GATGCCCTTGTCAATG	437	
Qy	121	AAAGGGACCCTGTGTTCATTTCCAGTACACAGAAGAAAATAATGTC-ATCCAATCGA	179	
Dd	438	AAAGGGCCCCCTGTGTTCATTTCCAGTACACAGAAGAAAATAATGTCAATCCGAATCGA	497	
Qy	180	GGGGTATATGCCCTTGGATTGCATGTTCTTGCAAGCCAGACCTTACACACATTCAGAAGACCT	239	
Dd	498	GGGGATATAGCCCTTGGATTGCATGTTCTTGCAAGCCAGACCTTACACACATTCAGAAGACCT	557	
Qy	240	TAAATAAACATGTCTTAAATGCAACACCGGCCTTACCTCTGTGAAACACGAGATTCTTTCGGGT	299	
Dd	558	TAAATAACATGTCTTAAATGCAACACCGGCCTTACCTCTGTGAAACGAGATTCTTTCGGGT	617	
Qy	300	TGAAGCAGAGTATCTCATGTCCGCTTGTGATAAAGTCAAGTCGAAACAGAACCTTCCCAAGA	359	

Db 618 TGAAGCAGAGTATCTCAGTCCTGTTATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 677
Qy 360 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGAGCAATTTAGAGTCGC 419
Db 678 AAAGAAATGCAAGGAAATGAATTTAGCTGTGAGTATGTGGGAGAGCAATTTAGAGTCGC 737
Qy 420 TTTTGTATGTTGAGATCCATGAGAACACAAAGATCTTTTCACTTACGGGTGTAAAT 479
Db 738 TTTTGTATGTTGAGATCCATGAGAACACAAAGATCTTTTCACTTACGGGTGTAAAT 797
Qy 480 GTGCGGAGAGAGMTTSSRSSAGCTTGGTTCTTTAAATAATCACATGCGGACACATATGG 539
Db 798 GTGCGGAGAGAGATTTCAAGGAGCTTGGTTCTTTAAATAATCACATGCGGACRATATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAACAGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAACAGATCAA 917
Qy 600 CGAGGTGCTCGAGGTGCACGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGT 659
Db 918 CGAGGTGCTCGAGGTGCACGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGT 977
Qy 660 TTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCACACAA 719
Db 978 TTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCACACAA 1037
Qy 720 AAAAAGCTGTTTTCGTACACAGCGGCGAGAGACTCTCCACAAAGGAGAAATGCCGTC 779
Db 1038 AAAAAGCTGTTTTCGTACACAGCGGCGAGAGACTCTCCACAAAGGAGAAATGCCGTC 1097
Qy 780 CTCGAGGAGGAGACTTCTCGAGTTGTTTCAAATTTGAGACCAAAATCTCACCTGAGAAAGGG 839
Db 1098 CTCGAGGAGGAGACTTCTCGAGTTGTTTCAAATTTGAGACCAAAATCTCACCTGAGAAAGGG 1157
Qy 840 GAAGAGCTGTCAGATGATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTGCCA 899
Db 1158 GAAGAGCTGTCAGATGATCCCTCAGCTCGATCCGTTTACCACTTCCAGGCTTGCCA 1217
Qy 900 GCTGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGAAATCGGGGCAAGAGG 959
Db 1218 KCTGCTACCAAGGAAAGTTGCCATTTGCCAAGAGTGAAGAAATTTGGGGCAAGAGG 1277
Qy 960 GAGCACGACAAACGACGATTCAGTTTCCGAGAGGAGCTTCGAGAAACAAATAAGGGCAG 1019
Db 1278 GAGCACGACAAACGACGATTCAGTTTCCGAGAGGAGCTTCGAGAAACAAATAAGAACCA 1337
Qy 1020 TTGTGAGGCTCTCGAAGAGAAAGAGAGTGCACACTCCACGCGAGCGCCCTC 1079
Db 1338 TTGTGAGGCTCTCGAAGAGAAAGAGAGTGCACACTCCACGCGAGCGCCCTC 1397
Qy 1080 CGTGACGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTG 1139
Db 1398 CGTGACGCGGATCCCAAGTTTACCCAGTAGCAAGGAGAGCCACTCACTGCTCCGAGTG 1457
Qy 1140 CGGCAAGCTTTTCAGAACTTACCACAGCTGGTCTTTCAGCTCCAGGGTCC 1189
Db 1458 CGGCAAGCTTTTCAGAACTTACCACAGCTGGTCTTTCAGCTCCAGGGTCC 1507

RESULT 9

US-08-892-695-3
; Sequence 3, Application US/08892695A
; Patent No. 6808878
; GENERAL INFORMATION:
; APPLICANT: Gray, Joe W
; APPLICANT: Collins, Collin
; APPLICANT: Hwang, Soo In
; APPLICANT: Godfrey, Tony
; APPLICANT: Kowel, David
; APPLICANT: Rommens, Johanna
; TITLE OF INVENTION: GENES FROM THE 20Q13 AMPLICON AND THEIR USES
; FILE REFERENCE: 2500.124US3
; CURRENT APPLICATION NUMBER: US/08/892,695A

; CURRENT FILING DATE: 1997-07-15
; EARLIER APPLICATION NUMBER: 08/785,532
; EARLIER FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/731,499
; EARLIER FILING DATE: 1996-10-16
; EARLIER APPLICATION NUMBER: 08/680,395
; EARLIER FILING DATE: 1996-07-15
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1507
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (447)
; OTHER INFORMATION: N is A, G, T, or U
US-08-892-695-3

Query Match 34.7%; Score 1106.2; DB 4; Length 1507;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1150; Conservative 9; Mismatches 28; Indels 3; Gaps 3;

Qy 1 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGATGG 60
Db 320 ATGCAATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGAT-GG 378
Qy 61 CCAGAGTGTGGCAGCTCTTTGGCAGTCCGATGAGATGGAGGATGCTTCTCAATG 120
Db 379 CCAAGAGTGTGGCAGCTCTTTGGCAGTCCGATGAGATGGGA-GATGCTTCTCAATG 437
Qy 121 AAAGGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTC-ATCCAAATCGA 179
Db 438 AAAGGGCCNCTGTTGTTCAATTCGAGCTACACAAGAAAAAATGTCATCCGATCGA 497
Qy 180 GGGGTATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 239
Db 498 GGGGAATATGCCCTTGGATTCATGTTCTGAGCCAGACCTTTCACATTCAGAAAGCCT 557
Qy 240 TAATAAATCATGCTTAATGCAACCGCCCTTACCTCTGTGAACGACGATCTTCCGGT 299
Db 558 TAATAAATCATGCTTAATGCAACCGCCCTTACCTCTGTGAACGACGATCTTCCGGT 617
Qy 300 TGAAGCAGAGTATCTCAGTCGCTTGAATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 359
Db 618 TGAAGCAGAGTATCTCAGTCGCTTGAATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGA 677
Qy 360 AAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGC 419
Db 678 AAAGAAATTCGAAGGAAATGAATTTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGC 737
Qy 420 TTTTGTATGTTGAGATCCATGAGAACACAAAGATCTTTTCACTTACGGGTGTAAAT 479
Db 738 TTTTGTATGTTGAGATCCATGAGAACACAAAGATCTTTTCACTTACGGGTGTAAAT 797
Qy 480 GTGCGGAGAGAGMTTSSRSSAGCTTGGTTCTTTAAATAATCACATGCGGACACATAATGG 539
Db 798 GTGCGGAGAGAGATTCAGGAGCTTGGTTCTTTAAATAATCACATGCGGACRATATGG 857
Qy 540 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAACAGATCAA 599
Db 858 CAAATCGGGGCCAGAGCAAACTGCAGCAAGGCTTGGAGAGTAGTCAGCAACAGATCAA 917
Qy 600 CGAGGTGCTCGAGGTGCACGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGT 659
Db 918 CGAGGTGCTCGAGGTGCACGCGCGGAGAGCATCTCTCTCTTACAAAATCTGCATGGT 977
Qy 660 TTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCACACAA 719
Db 978 TTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTAATTTGAGCAGCGCAAGGTGCACACAA 1037

QY 720 AAAAACTGTTTCGGTACACAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTC 779
Db 1038 AAAAACTGTTTCGGTACACAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTC 1097
QY 780 CTCGAGGAGGAGCTTCCTGCGAGTCTTCAACTTTGAGACCAAAATCTCACCCCTGAAACGGG 839
Db 1098 CTCGAGGAGGAGCTTCCTGCGAGTCTTCAACTTTGAGACCAAAATCTCACCCCTGAAACGGG 1157
QY 840 GAAGAAGCTGTGAGATGATCCCTCAGCTCGATCCGTTACCACTTCCAGGCTTTGGCA 899
Db 1158 GAAGAAGCTGTGAGATGATCCCTCAGCTCGATCCGTTACCACTTCCAGGCTTTGGCA 1217
QY 900 GCTGCTACCAAAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAAATCGGGCAGAGAGG 959
Db 1218 KCTGCTACCAAAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAAATCGGGCAGAGAGG 1277
QY 960 GAGCACCGACACGAGATTCAGTTCGAGAGGAGGCTTGAGAGAAACAAATTAAGGGCAG 1019
Db 1278 GAGCACCGACACGAGATTCAGTTCGAGAGGAGGCTTGAGAGAAACAAATTAAGAGCA 1337
QY 1020 TTGTGAGGCTCTCGAAGAGAAAGAGAGAGTGCAAAACACTCCCAAGCGGAGCGCCCTC 1079
Db 1338 TTGTGAGGCTCTCGAAGAGAAAGAGAGTGCAAAACACTCCCAAGCGGAGCGCCCTC 1397
QY 1080 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGAGAGGAGCCACTCACTGCTCCGAGTG 1139
Db 1398 CGTGACCGGATCCCAAGTTACCAAGTAGCAAGAGAGGAGCCACTCACTGCTCCGAGTG 1457
QY 1140 CGGAAAAGCTTTCAGAACTACCAAGCTGGTCTTGCACTCCAGGGTCC 1189
Db 1458 CGGAAAAGCTTTCAGAACTACCAAGCTGGTCTTGCACTCCAGGGTCC 1507

RESULT 10

US-09-949-016-34700/c
; Sequence 34700, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34700
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34700

Query Match 18.9%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 1.8e-175;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 726 TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTCCTCGAG 785
Db 601 TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTCCTCGAG 542
QY 786 GGAGACTTCTGCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGAGAA 845
Db 541 GGAGACTTCTGCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGAGAA 482
QY 846 GCCTGTGAGATGATCCCTCAGCTCGATCCGTTACCAAGCTTCCAGGCTTGGCAGTGGC 905
Db 481 GCCTGTGAGATGATCCCTCAGCTCGATCCGTTACCAAGCTTCCAGGCTTGGCAGTGGC 422

QY 906 TACCAAGGAAAAAGTTGCGCATTTGCCAAGAAAGTGAAGGAAATCGGGCAGAGGAGCAGC 965
Db 421 TACCAAGGAAAAAGTTGCGCATTTGCCAAGAAAGTGAAGGAAATCGGGCAGAGGAGCAGC 362
QY 966 CGAACACGACCATTCGAGTTCGGAGAGGAGCTTGAGAGAAACAAATTAAGGGCAGTTGTGC 1025
Db 361 CGAACACGACCATTCGAGTTCGGAGAGGAGCTTGAGAGAAACAAATTAAGGGCAGTTGTGC 302
QY 1026 AGGCTCTTCGCAAGAGAAAGAGTGCACACACTCCACGCGAGAGCGCCCTCCGTGGA 1085
Db 301 WGGCTCTTCGCAAGAGAAAGAGTGCACACACTCCACGCGAGAGCGCCCTCCGTGGA 242
QY 1086 CGCGATCCCAAGTTACCAAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGGGCAA 1145
Db 241 CGCGATCCCAAGTTACCAAGTAGCAAGAGAGAGCCCACTCACTGCTCCGAGTGGGCAA 182
QY 1146 AGCTTTTCAAGACCTTACCAAGTGGTCTTGCACCTCCAGGGTCCACAGAGGAGCGGAG 1205
Db 181 AGCTTTTCAAGACCTTACCAAGTGGTCTTGCACCTCCAGGGTCCACAGAGGAGCGGAG 122
QY 1206 GCGCGCGCGAGTCCGCCACCATGTCTGTGACGCGGAGGAGCGCGGAGAGCTGTCTCC 1265
Db 121 GCGCGCGCGAGTCCGCCACCATGTCTGTGACGCGGAGGAGCGCGGAGAGCTGTCTCC 62
QY 1266 TGACTCTCCGCCCTCTTGAGTGAATAATGAGAGCGCTGGATCGAGGGAGAGTGTGTTGGA 1325
Db 61 TGACTCTCCGCCCTCTTGAGTGAATAATGAGAGCGCTGGATCGAGGGAGAGTGTGTTGGA 2
QY 1326 A 1326
Db 1 A 1

RESULT 11

US-09-949-016-151880/c
; Sequence 151880, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151880
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151880

Query Match 18.9%; Score 600.6; DB 4; Length 601;
Best Local Similarity 99.8%; Pred. No. 1.8e-175;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 726 TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTCCTCGAG 785
Db 601 TGCTTTTCGGTACCAAGCGCGCAGACAGACTCTCCAAAGAGGAAATGCGGTCCTCGAG 542
QY 786 GGAGACTTCTGCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGAGAA 845
Db 541 GGAGACTTCTGCGAGTTGTTCAACTTGAGACCAAAATCTCACCCCTGAAACGGGAGAA 482
QY 846 GCCTGTGAGATGATCCCTCAGCTCGATCCGTTACCAAGCTTCCAGGCTTGGCAGTGGC 905

Db 481 GCCTGTGAGATGCATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGCGAGCTGGC 422
Qy 906 TACCAAGGAAAGTTGTCATTTGCGAAGAGTGAAGGAATCGGGGCAAGAGGAGCAC 965
Db 421 TACCAAGGAAAGTTGTCATTTGCGAAGAGTGAAGGAATCGGGGCAAGAGGAGCAC 362
Qy 966 CGAACAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAGGCGAGTTGTGC 1025
Db 361 CGAACAACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAAACAAATAGGCGAGTTGTGC 302
Qy 1026 AGGCTCTCGAAGAAAGAGTGAAGTCAAACTCCACCGGCAAGCGCCCTCGGTGGA 1085
Db 301 WGGCTCTCGAAGAAAGAGTGAAGTCAAACTCCACCGGCAAGCGCCCTCGGTGGA 242
Qy 1086 CGCGGATCCCAAGTTACCCAGTACGAGGAAGCACTCACTGCTCCGAGTGGCGAA 1145
Db 241 CGCGGATCCCAAGTTACCCAGTACGAGGAAGCACTCACTGCTCCGAGTGGCGAA 182
Qy 1146 AGCTTTCAAGAACCTACCAACAGCTGGTCTTGCACTTCCAGGGTCCACAGAGGACCGGAG 1205
Db 181 AGCTTTCAAGAACCTACCAACAGCTGGTCTTGCACTTCCAGGGTCCACAGAGGACCGGAG 122
Qy 1206 GCGCGGCGGAGTGCACCATGTCTGTGGAAGGAGGAGCGGGGACGTGTTCTCC 1265
Db 121 GCGCGGCGGAGTGCACCATGTCTGTGGAAGGAGGAGCGGGGACGTGTTCTCC 62
Qy 1266 TGACCTCGCCGCCCTCTGGATGAAGTGAAGCGTGGATCGAGGGAGGTTGTTCTGA 1325
Db 61 TGACCTCGCCGCCCTCTGGATGAAGTGAAGCGTGGATCGAGGGAGGTTGTTCTGA 2
Qy 1326 A 1326
Db 1 A 1

RESULT 12

US-09-949-016-34699/c
; Sequence 34699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-34699

Query Match 18.8%; Score 598.2; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 9.9e-175;
Matches 594; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Qy 99 GATGAGGATGCTTGTCAATGAAGGAGGACCGCTGTTTCCATTCGAGCTACACAAGA 158
Db 601 GATGAGGATGCTTGTCAATGAAGGAGGACCGCTGTTTCCATTCGAGCTACACAAGA 542
Qy 159 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 218
Db 541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482
Qy 219 CTTCAACATTCAGAAGACCTTAAATAACATGTCTTAAATGAACACCGGCTTACCCTCTG 278

Db 481 CTTCAACATTCAGAAGACCTTAAATAACATGTCTTAAATGAACACCGGCTTACCCTCTG 422
Qy 279 TGAACAGAGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 338
Db 421 TGAACAGAGCTTCTTCGGGTTGAAGCAGAGTATCTCAGTCCGCTTGATAAAAGTCAAGT 362
Qy 339 GCGAAACAGAACTCCCAAGGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATG 398
Db 361 GCGAAACAGAACTCCCAAGGAAAGAAATTCGAAGAAATGAATTTAGCTGTGAGGTATG 302
Qy 399 TGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCATGAGAACACACAAGATTC 458
Db 301 YGGGACAGACATTTAGAGTCGCTTTTGTGATGTTGATCCATGATGAAGAACACAACAAGATTC 242
Qy 459 TTTCACTTACGGGTGTAACATGTCGGAAGAGMTTSRRSAGCTTTGGTTTCTTAAAAA 518
Db 241 TTTCACTTACGGGTGTAACATGTCGGAAGAGMTTCAAGGAGCTTTGGTTTCTTAAAAA 182
Qy 519 TCACATCGGACACATATATGGCAATCGGGGGCCAGAACTGCAGCAAGGCTTGA 578
Db 181 TCACATCGGACACATATATGGCAATCGGGGGCCAGAACTGCAGCAAGGCTTGA 122
Qy 579 GAGTAGTCCAGAACGATCAACAGAGTCTCCAGGTGCAACGCGCCGAGAGCATCTCTC 638
Db 121 GAGTAGTCCAGAACGATCAACAGAGTCTCCAGGTGCAACGCGCCGAGAGCATCTCTC 62
Qy 639 TCCTTACAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGA 698
Db 61 TCCTTACAAATCTGCATGTTTGTGGCTTCTTATTTCCAAATAAAGAAAGTCTTAATTGA 2
Qy 699 G 699
Db 1 G 1

RESULT 13

US-09-949-016-151879/c
; Sequence 151879, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 151879
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-151879

Query Match 18.8%; Score 598.2; DB 4; Length 601;
Best Local Similarity 98.8%; Pred. No. 9.9e-175;
Matches 594; Conservative 7; Mismatches 0; Indels 0; Gaps 0;
Qy 99 GATGAGGATGCTTGTCAATGAAGGAGGACCGCTGTTTCCATTCGAGCTACACAAGA 158
Db 601 GATGAGGATGCTTGTCAATGAAGGAGGACCGCTGTTTCCATTCGAGCTACACAAGA 542
Qy 159 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 218
Db 541 AAAAAATGTCATCCAAATCGAGGGGTATATGCCCTTGGATTGCATGTTCTGCAGCCAGAC 482

```
QY 219 CTTACACATTCCAGAGACCTTTAATAAATGCTCTTAATGCAACACCGGCTTACCTCTG 278
Db 481 CTTACACATTCCAGAGACCTTTAATAAATGCTCTTAATGCAACACCGGCTTACCTCTG 422
QY 279 TGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGTATAAAGTCAAGT 338
Db 421 TGAACCCAGCAGTTCTTCGGGTTGAAGCAGAGTATCTCAGTCGCTTGTATAAAGTCAAGT 362
QY 339 GCGAACAGAACTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGTGTAGGTTATG 398
Db 361 GCGAACAGAACTCCCAAGGAAAGAAATGCAAGGAAATGAATTTAGTGTAGGTTATG 302
QY 399 TGGGCGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGCAACACACAAAGATT 458
Db 242 TGGGCGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCACATGCAACACACAAAGATT 424
QY 459 TTTCACTTACCGGTTGTAACATGTGCGGAAGAGMTTSSRRSAGCCTTGGTTTCTTAAAAA 518
Db 241 TTTCACTTACCGGTTGTAACATGTGCGGAAGAGATTCAAGGAGCCTTGGTTTCTTAAAAA 182
QY 519 TCACATGCGGACACATATGCGAAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGA 578
Db 181 TCACATGCGGACACATATGCGAAATCGGGGGCCAGAGCAAACTGCGAGCAAGGCTTGA 122
QY 579 GAGTAGTCCAGCAACGATCAACGAGTGTCCAGGTGCAACGCGCGGAGAGCATCTCTC 638
Db 121 GAGTAGTCCAGCAACGATCAACGAGTGTCCAGGTGCAACGCGCGGAGAGCATCTCTC 62
QY 639 TCCCTTACAAATCTCATGTTTGGGTTTGGCTTCTTATTTCCAAATAAGAAAGTCTAATGA 698
Db 61 TCCCTTACAAATCTCATGTTTGGGTTTGGCTTCTTATTTCCAAATAAGAAAGTCTAATGA 2
QY 699 G 699
Db 1 G 1
```

RESULT 14

US-09-620-312D-61

; Sequence 61, Application US/09620312D

; Patent No. 6569662

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Liu, Chenghua

; APPLICANT: Asundi, Vinod

; APPLICANT: Zhang, Jie

; APPLICANT: Ren, Feiyan

; APPLICANT: Chen, Rui-hong

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wehrman, Tom

; APPLICANT: Xue, Aidong J.

; APPLICANT: Yang, Yonghong

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yundong

; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei

; APPLICANT: John Tillinghast

; APPLICANT: Dmanac, Radoje T.

; TITLE OF INVENTION: No. 6569662el Nucleic Acids and

; TITLE OF INVENTION: Polypeptides

; FILE REFERENCE: 784CIP2B

; CURRENT APPLICATION NUMBER: US/09/620,312D

; CURRENT FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317

; PRIOR FILING DATE: 2000-04-25

; PRIOR APPLICATION NUMBER: 09/488,725

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1105

; SOFTWARE: pt_FL_genes Version 1.0

; SEQ ID NO 61

; LENGTH: 2765

; TYPE: DNA

;

```
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (123)..(2291)
US-09-620-312D-61
```

Query Match 2.1%; Score 66.8; DB 4; Length 2765;

Best Local Similarity 59.4%; Pred. No. 3e-09; Mismatches 6; Indels 0; Gaps 0;

Matches 101; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 379 GAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCAC 438

Db 939 GAGTTCCGCTGCCAAGTGTGCGGCGAGAGCTTTACACAGTCTTGTCTTCTCAAGGGCCAC 998

QY 439 ATGAGACACACAAAGATTCTTTCCTTACCGGTGTAAACATGTGCGGGAAGAGMTTSSRS 498

Db 999 ATGCGTAAGCACAAAGGCTCTTCGATCATGCGTGTCCGGTGTGCGGCGCTGCTTCAAG 1058

QY 499 SAGCCTTGGTTCTTAAAAATCACATGCGGACACATAATGCAATCGG 548

Db 1059 GAGCCCTGGTCTTTAAGAACACATGAAGGTGCACGCCAGCAAGCTGGG 1108

RESULT 15

US-09-949-016-4961

; Sequence 4961, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4961

; LENGTH: 2799

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-4961

Query Match 2.1%; Score 66.8; DB 4; Length 2799;

Best Local Similarity 59.4%; Pred. No. 3e-09; Mismatches 6; Indels 0; Gaps 0;

Matches 101; Conservative 6; Mismatches 63; Indels 0; Gaps 0;

QY 379 GAATTTAGCTGTGAGGTATGTGGGCGAGACATTTAGAGTCGCTTTTGTATGTTGAGATCCAC 438

Db 988 GAGTTCCGCTGCCAAGTGTGCGGCGAGAGCTTTACACAGTCTTGTCTTCTCAAGGGCCAC 1047

QY 439 ATGAGACACACAAAGATTCTTTCCTTACCGGTGTAAACATGTGCGGGAAGAGMTTSSRS 498

Db 1048 ATGCGTAAGCACAAAGGCTCTTCGATCATGCGTGTCCGGTGTGCGGCGCTGCTTCAAG 1107

QY 499 SAGCCTTGGTTCTTAAAAATCACATGCGGACACATAATGCAATCGG 548

Db 1108 GAGCCCTGGTCTTTAAGAACACATGAAGGTGCACGCCAGCAAGCTGGG 1157

Search completed: July 3, 2005, 16:59:39

Job time : 509.445 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 3, 2005, 16:40:17 ; Search time 1858.96 Seconds
(without alignments)
10749.639 Million cell updates/sec

Title: US-08-731-499-10
Perfect score: 3186
Sequence: 1 ATGCACTCGAAGTGACAGG.....GGAACACTACAGTGTGTAA 3186

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 6313374 seqs, 3136092125 residues

Total number of hits satisfying chosen parameters: 12626748

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

```

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3183.6	99.9	3186	US-08-731-499-10	Sequence 10, Appl
2	3183.6	99.9	5252	13 US-10-087-192-1709	Sequence 1709, Ap
3	3040.6	95.4	5632	15 US-10-177-293-505	Sequence 505, App
4	1561.2	49.0	36022	13 US-10-087-192-1708	Sequence 1708, Ap
5	1559.6	49.0	10365	8 US-08-731-499-9	Sequence 9, Appl
6	1511.2	47.4	3016	13 US-10-087-192-1706	Sequence 1706, Ap
7	1106.2	34.7	1507	8 US-08-731-499-3	Sequence 3, Appl

8	733	23.0	26345	13	US-10-087-192-1705	Sequence 1705, Ap
9	444.4	13.9	469	13	US-10-040-739-520	Sequence 520, App
c 10	401.4	12.6	530	16	US-10-029-386-9711	Sequence 9711, Ap
c 11	251	7.9	251	16	US-10-029-386-23411	Sequence 23411, A
12	164	5.1	267	9	US-09-783-590-9057	Sequence 9057, Ap
13	102.4	3.2	6033	17	US-10-172-118-1511	Sequence 1511, Ap
14	102.4	3.2	6033	18	US-10-342-887-1511	Sequence 1511, Ap
15	102.4	3.2	8156	14	US-10-074-475-93	Sequence 93, Appl
16	70	2.2	70	20	US-10-758-307-110	Sequence 110, Appl
17	70	2.2	70	21	US-10-852-797-372	Sequence 372, App
c 18	66.8	2.1	1014	16	US-10-029-386-20910	Sequence 20910, A
c 19	66.8	2.1	1229	16	US-10-029-386-20193	Sequence 20193, A
c 20	66.8	2.1	2765	15	US-10-037-270-61	Sequence 61, Appl
21	66.8	2.1	2765	17	US-10-117-722-61	Sequence 61, Appl
22	66.8	2.1	3039	17	US-10-172-118-1644	Sequence 1644, Ap
23	66.8	2.1	3039	18	US-10-342-887-1644	Sequence 1644, Ap
24	66.8	2.1	3617	9	US-09-764-847-1394	Sequence 1394, Ap
25	66.8	2.1	3617	9	US-09-764-847-1396	Sequence 1396, Ap
26	66.8	2.1	3617	14	US-10-092-154-1394	Sequence 1394, Ap
27	66.8	2.1	3617	14	US-10-092-154-1396	Sequence 1396, Ap
28	63.4	2.0	5154	21	US-10-956-157-1977	Sequence 1977, Ap
29	62.8	2.0	1033	16	US-10-029-386-22959	Sequence 22959, A
30	62.8	2.0	2918	17	US-10-104-047-835	Sequence 835, App
31	62.8	2.0	4599	10	US-09-764-872-787	Sequence 787, App
32	62.8	2.0	4599	10	US-09-764-872-790	Sequence 790, App
33	62.8	2.0	4603	10	US-09-764-872-791	Sequence 791, App
c 34	62.2	2.0	308	16	US-10-029-386-15037	Sequence 15037, A
c 35	62.2	2.0	592	16	US-10-029-386-1335	Sequence 1335, Ap
c 36	61.2	1.9	193691	20	US-10-719-993-6768	Sequence 6768, Ap
c 37	61.2	1.9	318760	20	US-10-719-993-6765	Sequence 6765, Ap
38	60.6	1.9	3257	21	US-10-956-157-1071	Sequence 1071, Ap
39	60	1.9	60	10	US-09-308-975-12266	Sequence 12266, A
40	59.8	1.9	453	10	US-09-918-995-5364	Sequence 5364, App
41	59.8	1.9	2703	17	US-10-264-049-841	Sequence 841, App
42	59.8	1.9	2785	17	US-10-120-988-105	Sequence 105, App
43	59.8	1.9	2901	9	US-09-974-298-117	Sequence 117, App
44	59.8	1.9	3120	17	US-10-094-749-337	Sequence 337, App
45	59.8	1.9	3646	10	US-09-814-353-20524	Sequence 20524, A

ALIGNMENTS

RESULT 1

US-08-731-499-10
; Sequence 10, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435


```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3186 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: - 1..3186
; LOCATION: 1..3186
; OTHER INFORMATION: /note= "ZABCl Open Reading Frame"
US-08-731-499-10

Query Match          99.9%; Score 3183.6; DB 8; Length 3186;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCNATCGAAGTACAGGAAACATGCCAACTCAATCCCTTTAATGTACATGGATGGG 60
Db 1 ATGCAATCGAAGTGACAGGAACATGCCAACTCAATCCCTTTAATGTACATGGATGGG 60

Qy 61 CCAGAAGTGATTGGCAGCTCTCTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
Db 61 CCAGAAGTGATTGGCAGCTCTCTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120

Qy 121 AAAGGACCGGTGTTGTTCCATTCCGAGCTACAAAGAAACAAATGTCCATCAATCGAG 180
Db 121 AAAGGACCGGTGTTGTTCCATTCCGAGCTACAAAGAAACAAATGTCCATCAATCGAG 180

Qy 181 GGGTATATGCCCTTGGATTGATGTTCTGACGACGACCTTACACATTCAGAGACCTT 240
Db 181 GGGTATATGCCCTTGGATTGATGTTCTGACGACGACCTTACACATTCAGAGACCTT 240

Qy 241 AATAAACATGCTTTAATCAACACCGGCTACCTCTCTGAAACAGCAGTCTCTCGGGTT 300
Db 241 AATAAACATGCTTTAATCAACACCGGCTACCTCTCTGAAACAGCAGTCTCTCGGGTT 300

Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGGAGACGAAACCTCCCAAGGAA 360
Db 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAGTCAAGTGGAGACGAAACCTCCCAAGGAA 360

Qy 361 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGATCGCT 420
Db 361 AAGAAATTCAGGAAATGAATTTAGCTGTGAGGTATGTGGGCAGACATTTAGATCGCT 420

Qy 421 TTTGATGTTGAGATCCATGAGAACACAAAGATTTCTTCACTTACGGGTGTAAATG 480
Db 421 TTTGATGTTGAGATCCATGAGAACACAAAGATTTCTTCACTTACGGGTGTAAATG 480

Qy 481 TCGGAAGAAGMTTSSRRSSAGCCTTGGTTTCTTAAAAATCATCGCGACACATAATGGC 540
Db 481 TCGGAAGAAGMTTSSRRSSAGCCTTGGTTTCTTAAAAATCATCGCGACACATAATGGC 540

Qy 541 AAATCGGGGGCCAGAGCAAACTGACGAGCTTGGAGATAGTCCAGCAACGATCAAC 600
Db 541 AAATCGGGGGCCAGAGCAAACTGACGAGCTTGGAGATAGTCCAGCAACGATCAAC 600

Qy 601 GAGGTCGTCCAGTGCACGCGGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 660
Db 601 GAGGTCGTCCAGTGCACGCGGCCGAGAGCATCTCTCTCTTACAAAATCTGATGGTT 660

Qy 661 TGTGCTTCTTATTTCCAAATAAAGAAAGTCTAAATTGAGACCGCAAGGTGCAACCAA 720
Db 661 TGTGCTTCTTATTTCCAAATAAAGAAAGTCTAAATTGAGACCGCAAGGTGCAACCAA 720
```

```
Db 661 TGTGCTTCTTATTTCCAAATAAAGAAAGTCTAAATTGAGACCGCAAGGTGCAACCAA 720
Qy 721 AAAACTGCTTTCGGTATCCAGCAGCGCGCAGACAGACTCTCCACAAGAGGAATGCGCTCC 780
Db 721 AAAACTGCTTTCGGTATCCAGCAGCGCGCAGACAGACTCTCCACAAGAGGAATGCGCTCC 780
Qy 781 TCGAGGAGGAGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCACTTGAACGGGG 840
Db 781 TCGAGGAGGAGACTTCTCGCAGTTGTTCAACTTGAGACCAAAATCTCACTTGAACGGGG 840
Qy 841 AAGAAAGCTGTGATGATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900
Db 841 AAGAAAGCTGTGATGATCCCTCAGCTCGATCCGTTCCACCACTTCCAGGCTTGGCAG 900
Qy 901 CTGGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAGGG 960
Db 901 CTGGCTACCAAGAAAGTTGCCATTTGCCAAGAGTGAAGGAATCGGGCAAGAAGGG 960
Qy 961 AGCACCGACAAACGAGATTCCGAGTTCGAGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT 1020
Db 961 AGCACCGACAAACGAGATTCCGAGTTCGAGAGAGGAGCTTGGAGAAACAAATAAGGGCAGT 1020
Qy 1021 TGTGAGGCTCTCGCAAGAGAAAGAGTGCACCACTCCACGGGGAAGCCCTCC 1080
Db 1021 TGTGAGGCTCTCGCAAGAGAAAGAGTGCACCACTCCACGGGGAAGCCCTCC 1080
Qy 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGAAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Db 1081 GTGACGCGGATCCCAAGTTTACCCAGTAGAAGAGAGCCCACTCACTGCTCCGAGTGC 1140
Qy 1141 GGCAAAAGCTTTTCCAGAACTTACCACAGCTGTGCACTCCAGGCTCCACAAGAGGAC 1200
Db 1141 GGCAAAAGCTTTTCCAGAACTTACCACAGCTGTGCACTCCAGGCTCCACAAGAGGAC 1200
Qy 1201 CGGAGGCGCGCGCGGAGTCCGCCACCATCTGTGAGCGGAGGAGCGCGGAGCGT 1260
Db 1201 CGGAGGCGCGCGCGGAGTCCGCCACCATCTGTGAGCGGAGGAGCGCGGAGCGT 1260
Qy 1261 TCTCTGACCTCGCGCGCCCTCTCGATGAAATGGAGCGGTGGATCGAGGGGAAGGTGGT 1320
Db 1261 TCTCTGACCTCGCGCGCCCTCTCGATGAAATGGAGCGGTGGATCGAGGGGAAGGTGGT 1320
Qy 1321 TCTGAAGACGGATCTGAGGATGGCTTCCGAAAGAAATCCATCTGGATATAAATGATGAT 1380
Db 1321 TCTGAAGACGGATCTGAGGATGGCTTCCGAAAGAAATCCATCTGGATATAAATGATGAT 1380
Qy 1381 GGAGGAAATAAACAATCTTACATCTTCAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1440
Db 1381 GGAGGAAATAAACAATCTTACATCTTCAGAGAGTGTAGTTATTTGGGAAAGTTTTTC 1440
Qy 1441 CGTTCAAAATTTACCTCAATATTCATCTCAGAACCGATACAGGTGAAAAAACCATACAAA 1500
Db 1441 CGTTCAAAATTTACCTCAATATTCATCTCAGAACCGATACAGGTGAAAAAACCATACAAA 1500
Qy 1501 TGTGAATTTTGTGAATATGTGCGAGCCCGAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Db 1501 TGTGAATTTTGTGAATATGTGCGAGCCCGAGAGACATCTCTGAGGTATCATCTTGGAGAGA 1560
Qy 1561 CATCAAGGAAATAAACAATCTTGTCTGAGTCAAGAGCGATGTTAAAAATCAG 1620
Db 1561 CATCAAGGAAATAAACAATCTTGTCTGAGTCAAGAGCGATGTTAAAAATCAG 1620
Qy 1621 GACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAACCAAAAAATTTGAAAAAGTTT 1680
Db 1621 GACACTGAAGATGCACTATTAAACCGCTGACAGTGCAGAAACCAAAAAATTTGAAAAAGTTT 1680
Qy 1681 TTTGATGTTGCCAAGATGTTTACAGGAGTCCACTCTGCAAGAGCTTAAAGAGATGCGCT 1740
Db 1681 TTTGATGTTGCCAAGATGTTTACAGGAGTCCACTCTGCAAGAGCTTAAAGAGATGCGCT 1740
Qy 1741 TCTGTTTTTCAAGATGTTCTCGGCGAGCGCTCTCTCTCACCAGCACAAAGATACTCAG 1800
Db 1741 TCTGTTTTTCAAGATGTTCTCGGCGAGCGCTCTCTCTCTCACCAGCACAAAGATACTCAG 1800
```

```
QY 1801 GATTTCCTAAATAAGTGCAGTCTGATGACAGTGTGATAAGTGAATAAATAAACCTTACCCCT 1860
DB 1801 GATTTCCTAAATAAGTGCAGTCTGATGACAGTGTGATAAGTGAATAAATAAACCTTACCCCT 1860
QY 1861 GCTTACCTGGAAGTGTAAATAAGAGATCAGCAGTGTGAATCTCAGGCAATAAATACCTCATC 1920
DB 1861 GCTTACCTGGAAGTGTAAATAAGAGATCAGCAGTGTGAATCTCAGGCAATAAATACCTCATC 1920
QY 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCAACCCATAAATACCTTGAA 1980
DB 1921 TGTAGAACCAAGCGGATGTTACTCTCTCCGGATGGCAGTACCAACCCATAAATACCTTGAA 1980
QY 1981 GTTAGCCCCAAGAGAGACAAAGCGAGACCGCAGTGTGATGAGATACAGGCAAGTGTG 2040
DB 1981 GTTAGCCCCAAGAGAGACAAAGCGAGACCGCAGTGTGATGAGATACAGGCAAGTGTG 2040
QY 2041 GATTGTACGAAATAACCTTTAAATTTATCCGTGGGGCTCTTCAATATGGCCCGCAAT 2100
DB 2041 GATTGTACGAAATAACCTTTAAATTTATCCGTGGGGCTCTTCAATATGGCCCGCAAT 2100
QY 2101 TCTTTGAGTAAAGTGTGATCCAAAGTATCACCTGTCCATTTGTACCTTCAAGCATTT 2160
DB 2101 TCTTTGAGTAAAGTGTGATCCAAAGTATCACCTGTCCATTTGTACCTTCAAGCATTT 2160
QY 2161 TATCCAGAGTATTAATGATGACACAGAGACTGGAGCATTAATACATCTCGACGTTAT 2220
DB 2161 TATCCAGAGTATTAATGATGACACAGAGACTGGAGCATTAATACATCTCGACGTTAT 2220
QY 2221 AAAAATGTGAAACCAAGTCTTGTGTAGAGTCCAGCTGACCGATGCCCGCAGCGTTG 2280
DB 2221 AAAAATGTGAAACCAAGTCTTGTGTAGAGTCCAGCTGACCGATGCCCGCAGCGTTG 2280
QY 2281 CTGGAAAGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2340
DB 2281 CTGGAAAGATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2340
QY 2341 CCGGCGCAGTCAATATCCCTGCTGATGCGAAGGGGAGCAGAGCCCTCTCTGGGCGAGG 2400
DB 2341 CCGGCGCAGTCAATATCCCTGCTGATGCGAAGGGGAGCAGAGCCCTCTCTGGGCGAGG 2400
QY 2401 AAGGCCCTCTGACTTTCAGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2460
DB 2401 AAGGCCCTCTGACTTTCAGGATAGACTCTAGCACTTTAGCCCCCAAGTAACCTGAAGTCC 2460
QY 2461 CACAGACCAACAGCAGATGTGGGGTCCAGGGGCGCCACACAGGCAACAGCAATCTGAG 2520
DB 2461 CACAGACCAACAGCAGATGTGGGGTCCAGGGGCGCCACACAGGCAACAGCAATCTGAG 2520
QY 2521 ATGTTTCTTAAACAGTGTGTTCCCTGACCGGATAGACAAATAAGACCCGAGCAAAA 2580
DB 2521 ATGTTTCTTAAACAGTGTGTTCCCTGACCGGATAGACAAATAAGACCCGAGCAAAA 2580
QY 2581 TTGAACCTCTTCCAGTAGTCTCTCTCTGAGCCACCTCGGCGAGCAGTAACATCAATGGT 2640
DB 2581 TTGAACCTCTTCCAGTAGTCTCTCTCTGAGCCACCTCGGCGAGCAGTAACATCAATGGT 2640
QY 2641 TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGACCTCTCCGGGAGAGACTATTTC 2700
DB 2641 TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGACCTCTCCGGGAGAGACTATTTC 2700
QY 2701 TGTAAATCGGAGTCCAGCAATATCTGAGCAGAAATTTGGTGAGCCCTTCCAAAAAGACTG 2760
DB 2701 TGTAAATCGGAGTCCAGCAATATCTGAGCAGAAATTTGGTGAGCCCTTCCAAAAAGACTG 2760
QY 2761 AGTCCAGGTGTTGCTTACCTTGACCTGACAGCCCGGGGCAATTAAGAGAGGCTAT 2820
DB 2761 AGTCCAGGTGTTGCTTACCTTGACCTGACAGCCCGGGGCAATTAAGAGAGGCTAT 2820
QY 2821 GACCTTCCCAAGTACATATGCTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG 2880
DB 2821 GACCTTCCCAAGTACATATGCTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG 2880
```

```
QY 2881 TATCGTGCAGGGCTGCTCCCAACCAAGGTTCTTGAGCTCCAGCGAGTTCGATTTCT 2940
DB 2881 TATCGTGCAGGGCTGCTCCCAACCAAGGTTCTTGAGCTCCAGCGAGTTCGATTTCT 2940
QY 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCCGGGCCACTTTTACACTTGTGTG 3000
DB 2941 CCAATGTGCTGACTGTTTCAAGAGCCCTATGCTGCTCCGGGCCACTTTTACACTTGTGTG 3000
QY 3001 CTTGCTGTGTAGTCCAGCATCCAGCTCGAGCTTTAGAGGTTTGTGTGGATGTCACTGCTTA 3060
DB 3001 CTTGCTGTGTAGTCCAGCATCCAGCTCGAGCTTTAGAGGTTTGTGTGGATGTCACTGCTTA 3060
QY 3061 CTCCCCATGAATTAATTTTACTTCTTCTTGAAGCGGATCGTGAAGCTACTGAA 3120
DB 3061 CTCCCCATGAATTAATTTTACTTCTTCTTGAAGCGGATCGTGAAGCTACTGAA 3120
QY 3121 ATAAGCTGTGATGTGATGATCAATAAATATATAGGATCTGCAAGGAACAATACAGTT 3180
DB 3121 ATAAGCTGTGATGTGATGATCAATAAATATATAGGATCTGCAAGGAACAATACAGTT 3180
QY 3181 GTGTAA 3186
DB 3181 GTGTAA 3186
```

RESULT 2

```
US-10-087-192-1709
; Sequence 1709, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1709
; LENGTH: 5252
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1709
```

```
Query Match 99.9%; Score 3183.6; DB 13; Length 5252;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 3180; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 ATGCAATCGAAGTGCAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGG 60
DB 25 ATGCAATCGAAGTGCAGGAAACATGCCAACTCAATCCCTTTAATGTACATGATGG 84
QY 61 CCAAGTGTGAGTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 120
DB 85 CCAAGTGTGAGTGGCAGCTCTTTGGCAGTCCGATGGAGATGGAGATGCCCTTGTCAATG 144
QY 121 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTATCAAAATCGAG 180
DB 145 AAAGGACCGCTGTTGTTCCATTCGAGCTACACAAGAAAAAATGTATCAAAATCGAG 204
QY 181 GGGTATATGCTTGGATTTGATGTCCTTTCAGCCAGACCTTCACATTCAGAACCTT 240
DB 205 GGGTATATGCTTGGATTTGATGTCCTTTCAGCCAGACCTTCACATTCAGAACCTT 264
QY 241 AATAAATGCTTAAATGCAACACCGGCTTACCTCTGTGAACACGAGCTTCTTCGGGT 300
DB 265 AATAAATGCTTAAATGCAACACCGGCTTACCTCTGTGAACACGAGCTTCTTCGGGT 324
```

QY 301 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAAGTGCAGAACAGAACCTCCCAAGGAA 360
Db 325 GAAGCAGAGTATCTCAGTCCGCTTGATATAAAGTCAAAGTGCAGAACAGAACCTCCCAAGGAA 384
QY 361 AAGAAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTGCCT 420
Db 385 AAGAAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGCGACATTTAGAGTGCCT 444
QY 421 TTTGATGTGAGATCCCATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 480
Db 445 TTTGATGTGAGATCCCATCAGAACACACAAAGATTTCTTCACTTACGGGTGTAAACATG 504
QY 481 TCGGGAAGAGTTSRRSSAGCTTGGTTTCTTAAATAATCATCGCGACACATAATGGC 540
Db 505 TCGGGAAGAGTTCAGAGGCTTGGTTTCTTAAATATCATGCGGACACATAATGGC 564
QY 541 AAATCGGGGGCAGAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 600
Db 565 AAATCGGGGGCAGAGCAAACTGCAGCAAGCTTTGGAGAGTAGTCCAGCAACGATCAAC 624
QY 601 GAGGTGCTCAGGTGCACGCGCGGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 660
Db 625 GAGGTGCTCAGGTGCACGCGCGGAGAGCATCTCCTCTCTTACAAATCTGCATGGTT 684
QY 661 TGTGCTTCTTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 720
Db 685 TGTGCTTCTTATTTCCAAATANAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAA 744
QY 721 AAAACTGCTTTCCGTAC CAGCAGCGCGCAGACAGACTCTCCACAGAGGAAATGCGGTCC 780
Db 745 AAAACTGCTTTCCGTAC CAGCAGCGCGCAGACAGACTCTCCACAGAGGAAATGCGGTCC 804
QY 781 TCGAGGAGGACTTCTCCTCAGTGTGTTCAACTGAGACCAAAATCTCACCTCTGAAACGGG 840
Db 805 TCGAGGAGGACTTCTCCTCAGTGTGTTCAACTGAGACCAAAATCTCACCTCTGAAACGGG 864
QY 841 AAGAACCTGT CAGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGCTTTGGCAG 900
Db 865 AAGAACCTGT CAGATGATCCCTCAGCTCGATCCGTTTCAACACCTTCCAGCTTTGGCAG 924
QY 901 CTGGCTACCAAAGGAAAGTTGCCATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAAGG 960
Db 925 CTGGCTACCAAAGGAAAGTTGCCATTTGCCAAGAAAGTGAAGGAATCGGGGCAAGAAGG 984
QY 961 AGCACGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAACAAATAAGGCGAGT 1020
Db 985 AGCACGACACGACGATTCGAGTTCCGAGAGGAGCTTGGAGAACAAATAAGGCGAGT 1044
QY 1021 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGCGGAGCGCCCTCC 1080
Db 1045 TGTGAGGCTCTCGCAAGAGAAAGAGAGTGCACCACTCCACGCGGAGCGCCCTCC 1104
QY 1081 GTGGAACGGATCCCAAGTTACCAAGTAGAAGGAGAGCCACCTCACTGCTCCGAGTGC 1140
Db 1105 GTGGAACGGATCCCAAGTTACCAAGTAGAAGGAGAGCCACCTCACTGCTCCGAGTGC 1164
QY 1141 GCGAAGCTTTCAGAACCTTACACACGCTGCTTCCACTCCAGGCTCCACAGAGGAC 1200
Db 1165 GCGAAGCTTTCAGAACCTTACACACGCTGCTTCCACTCCAGGCTCCACAGAGGAC 1224
QY 1201 CGGAGGCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGCAGCGGGGACGTGT 1260
Db 1225 CGGAGGCGGCGGAGTCCGCCACCATGTCTGTGACGGGAGCAGCGGGGACGTGT 1284
QY 1261 TCTCTGACCTCCGCCCTCTGATGAAATGAGCGCTGGATCGAGGGGAAAGGTGT 1320
Db 1285 TCTCTGACCTCCGCCCTCTGATGAAATGAGCGCTGGATCGAGGGGAAAGGTGT 1344
QY 1321 TCTGAGACGATCTGAGATGGCTTCCGAGGAAATCCATCTCGATAAATAATGATAT 1380
Db 1345 TCTGAGACGATCTGAGATGGCTTCCGAGGAAATCCATCTCGATAAATAATGATAT 1404
QY 1381 GGAGGAAAAATAAACATCTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAGTTTTTC 1440

Db 1405 GGAGGAAAAATAAACATCTTACATCTTCAAGAGGTGTAGTTATTGTGGAAGTTTTTC 1464
QY 1441 CGTTCAAAATTTATACCTCAATATTTATCTCAGAACGCATACAGGTGAAACACCATACAAA 1500
Db 1465 CGTTCAAAATTTATACCTCAATATTTATCTCAGAACGCATACAGGTGAAACACCATACAAA 1524
QY 1501 TGTGAAATTTTGTGAATATGTCGAGCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1560
Db 1525 TGTGAAATTTTGTGAATATGTCGAGCCGAGAACATCTCTGAGGTATCACTTGGAGAGA 1584
QY 1561 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1620
Db 1585 CATCAAGGAAAAACAAACCGATGTTGCTGCTGAAGTCAAGAACGATGGTAAATAATCAG 1644
QY 1621 GACACTCAAGATGACATTTAAACCGCTGACAGTGGCAACCCAAATAATTTGAAAGATTT 1680
Db 1645 GACACTCAAGATGACATTTAAACCGCTGACAGTGGCAACCCAAATAATTTGAAAGATTT 1704
QY 1681 TTTGATGTGCCAAAGATGTTACAGGAGTCCACCTGCAAGCAGCTTTAAGGAGATGCT 1740
Db 1705 TTTGATGTGCCAAAGATGTTACAGGAGTCCACCTGCAAGCAGCTTTAAGGAGATGCT 1764
QY 1741 TCTGTTTTTCAGAAATGTTTGGGCGAGCGTCTCTCTCACAGCACACAAAGATCTCAG 1800
Db 1765 TCTGTTTTTCAGAAATGTTTGGGCGAGCGTCTCTCTCACAGCACACAAAGATCTCAG 1824
QY 1801 GATTTCCATAAATAATGAGCTGATGACAGTCTCATAAAGTGAATAAATAATTTTACCCCT 1860
Db 1825 GATTTCCATAAATAATGAGCTGATGACAGTCTCATAAAGTGAATAAATAATTTTACCCCT 1884
QY 1861 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAATAAATCCTCATC 1920
Db 1885 GCTTACCTGGACCTGTTTAAAAAAGAGATCAGCAGTTGAAACTCAGGCAATAAATCCTCATC 1944
QY 1921 TGTAGAACCAAGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATAAATCCTTGA 1980
Db 1945 TGTAGAACCAAGGCGGATGTTTACTCTCTCCGATGGCAGTACACCCATAAATCCTTGA 2004
QY 1981 GTTAGCCCCCAAGAGAACCAAGGAGACCGACGCTGACATCAGATACAGGCCAGTGTG 2040
Db 2005 GTTAGCCCCCAAGAGAACCAAGGAGACCGACGCTGACATCAGATACAGGCCAGTGTG 2064
QY 2041 GATTTGTCACGAAAAACCTTTTAAATTTATFCCGTGGGGCTTTTCAAAATTTGCCGCAAT 2100
Db 2065 GATTTGTCACGAAAAACCTTTTAAATTTATFCCGTGGGGCTTTTCAAAATTTGCCGCAAT 2124
QY 2101 TCTTTGAGTAAAAAGTTTGAATTCAGATATCACCTGTCCATTTTGTACCTTCAAGACATTT 2160
Db 2125 TCTTTGAGTAAAAAGTTTGAATTCAGATATCACCTGTCCATTTTGTACCTTCAAGACATTT 2184
QY 2161 TATCCAGAGTAAATAATGATGACACAGAGCTGGAGCATAAATACAACTCTGACGTTTCA 2220
Db 2185 TATCCAGAGTAAATAATGATGACACAGAGCTGGAGCATAAATACAACTCTGACGTTTCA 2244
QY 2221 AAAAATCTGCAAAACAACTCTTGTAGAAAGTGCACGTACCGGATGCCGCGCAGCTTG 2280
Db 2245 AAAAATCTGCAAAACAACTCTTGTAGAAAGTGCACGTACCGGATGCCGCGCAGCTTG 2304
QY 2281 CTGGGAAAGATGTGCTTCCCTCTCTAGTTTCTGTAAACCCAGCCCAAGTCTGCTTTC 2340
Db 2305 CTGGGAAAGATGTGCTTCCCTCTCTAGTTTCTGTAAACCCAGCCCAAGTCTGCTTTC 2364
QY 2341 CGGCGCAGTCCAAATCCCTGCCATCTCGAAGGGGAGAGAGCCCTCTTGGGGCAGGC 2400
Db 2365 CGGCGCAGTCCAAATCCCTGCCATCTCGAAGGGGAGAGAGCCCTCTTGGGGCAGGC 2424
QY 2401 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTTAGCCCCCAAGTAACTGAAGTCC 2460
Db 2425 AAGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTTAGCCCCCAAGTAACTGAAGTCC 2484
QY 2461 CACAGACACAGCAGAAATGTGGGGTCCAAAGGGCGGCCACAGGCCAACAGCAATCTGAG 2520

Db 2485 CACAGACCACAGCAGAAATGTGGGGTCCAAAGGGCGCCACACAGGCAACAGCAATCTGAG 2544
Qy 2521 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGTAAGACAAAAGACCCGAGACAAA 2580
Db 2545 ATGTTTCTTAAACCAAGTGTTCCTCTGACCGGTAAGACAAAAGACCCGAGACAAA 2604
Qy 2581 TTGAACCTCTTCCAGTAGCTCTTCTCAGCCACCCCTCGGACAGTAACATCAATGCT 2640
Db 2605 TTGAACCTCTTCCAGTAGCTCTTCTCAGCCACCCCTCGGACAGTAACATCAATGCT 2664
Qy 2641 TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACTTCCCGGAAGAGACTATTTC 2700
Db 2665 TCCATCGACTACCCCGCAAGAACGACAGCCCGTGGGCACTTCCCGGAAGAGACTATTTC 2724
Qy 2701 TGTATCGAGTCCAGCAGTAATCTGACAGCAGTAATTTGTGAGCCCTTCCAAAAGACTG 2760
Db 2725 TGTATCGAGTCCAGCAGTAATCTGACAGCAGTAATTTGTGAGCCCTTCCAAAAGACTG 2784
Qy 2761 AAGTCCAGGTGTGGTCCCTTCCAGTGTGACAGCCCGGGCCAAATTACAGAGAGGCTAT 2820
Db 2785 AAGTCCAGGTGTGGTCCCTTCCAGTGTGACAGCCCGGGCCAAATTACAGAGAGGCTAT 2844
Qy 2821 GACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG 2880
Db 2845 GACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACTGTTACCGCAGGACTGTGTG 2904
Qy 2881 TATCGTGCAGCGCTGCTCCCAACCAAGGTTCTGAGTCTCAGGAGGTCGATTC 2940
Db 2905 TATCGTGCAGCGCTGCTCCCAACCAAGGTTCTGAGTCTCAGGAGGTCGATTC 2964
Qy 2941 CCAATGTGCTGACTGTTCCAGAGCCCTATGTTGCTCGGCGCACTTTACACTTGTGTG 3000
Db 2965 CCAATGTGCTGACTGTTCCAGAGCCCTATGTTGCTCGGCGCACTTTACACTTGTGTG 3024
Qy 3001 CTGCTGTGAGTCCAGAGTCCAGTCCAGTTCAGAGGTCCTGAGGTCCTGAGTGTG 3060
Db 3025 CTGCTGTGAGTCCAGAGTCCAGTTCAGAGGTCCTGAGGTCCTGAGTGTG 3084
Qy 3061 CTCCCATGAATTAATTTTACTTCATCTCTTTGAGAGCGAATGTTGAAAGCTACTGAA 3120
Db 3085 CTCCCATGAATTAATTTTACTTCATCTCTTTGAGAGCGAATGTTGAAAGCTACTGAA 3144
Qy 3121 ATAAGCTGTGATGTACTGTACATATAAATATAGGAATCTGCAAGGACACTACGTT 3180
Db 3145 ATAAGCTGTGATGTACTGTACATATAAATATAGGAATCTGCAAGGACACTACGTT 3204
Qy 3181 GTGTAA 3186
Db 3205 GTGTAA 3210

RESULT 3

US-10-177-293-505
; Sequence 505, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatt, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Gannavarpu, Manjula
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul

; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-038
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 505
; LENGTH: 5632
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-177-293-505

Query Match 95.4%; Score 3040.6; DB 15; Length 5632;

Best Local Similarity 95.8%; Pred. No. 0;

Matches 3180; Conservative 6; Mismatches 0; Indels 133; Gaps 1;

Qy 1 ATGCATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGATGGG 60
Db 272 ATGCATCGAAAGTGACAGGAAACATGCCAACTCAATCCCTCTTAAATGTACATGATGGG 331
Qy 61 CCAGAAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGSCCTTTGTCAATG 120
Db 332 CCAGAAGTGATTTGGCAGCTCTCTTGGCAGTCCGATGGAGATGGAGATGSCCTTTGTCAATG 391
Qy 121 AAAGGACCGCTGTTGTTTCCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 180
Db 392 AAAGGACCGCTGTTGTTTCCATTCGAGCTACACAAGAAAAAATGTTCATCCAAATCGAG 451
Qy 181 GGGTATATGCCCTTGGATTCATGTTCTGCAGCCAGACCTTCACACATTTCAGAACCTT 240
Db 452 GGGTATATGCCCTTGGATTCATGTTCTGCAGCCAGACCTTCACACATTTCAGAACCTT 511
Qy 241 AATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTGAACACGAGTTCCTCGGGTT 300
Db 512 AATAAACATGTTTAAATGCAACACCGGCTTACCCTCTGTGAACACGAGTTCCTCGGGTT 571
Qy 301 GAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGTCCGAAACAGAACCTTCCCAAGGAA 360
Db 572 GAAGCAGAGTATCTCAGTCCGCTTGTATATAAAGTCAAGTCCGAAACAGAACCTTCCCAAGGAA 631
Qy 361 AAGAAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTAGATCGCT 420
Db 632 AAGAAATTCAGAGGAAATGAATTTAGCTGTGAGGTATGTGGGACAGACATTAGATCGCT 691
Qy 421 TTTGATGTTGAGATCCACATGAGAACACACAAAAGATCTTTTACCTTACGGGTGTAACATG 480
Db 692 TTTGATGTTGAGATCCACATGAGAACACACAAAAGATCTTTTACCTTACGGGTGTAACATG 751
Qy 481 TCGGAAAGAGMTTSSRSSAGCCTTGGTTTCTTAAAAATTCATCGGACACATATGCG 540
Db 752 TCGGAAAGAGMTTSSRSSAGCCTTGGTTTCTTAAAAATTCATCGGACACATATGCG 811
Qy 541 AAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTTGGAGATGTATCCAGCAACGATCAAC 600
Db 812 AAATCGGGGGCCAGAGCAAACTGCAGCAAGGCTTTGGAGATGTATCCAGCAACGATCAAC 871
Qy 601 GAGTCTGTCCAGGTGACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGTT 660
Db 872 GAGTCTGTCCAGGTGACGCGCCGAGAGCATCTCTCTCTTACAAAATCTGCATGTT 931

Qy	661	TGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCAACAAA	720
Db	932	TGTGGCTTCTCTATTTCCAAATAAAGAAAGTCTAATTGAGCACCGCAAGGTGCAACAAA	991
Qy	721	AAAATCTGTTTCGTATCACGAGCGCGACAGACTCTCCACAAGGAGAAATGCCGTCC	780
Db	992	AAAATCTGTTTCGTATCACGAGCGCGACAGACTCTCCACAAGGAGAAATGCCGTCC	1051
Qy	781	TCGAGGGAGGACTTCCTGCAGTTGTTTCAAATTTGAGACAAAATCTCACCTCGAAACGGGG	840
Db	1052	TCGAGGGAGGACTTCCTGCAGTTGTTTCAAATTTGAGACAAAATCTCACCTCGAAACGGGG	1111
Qy	841	AAGAAGCCTGTTCAGATGCATCCCTCAGTTCGATCCGTTTCACCTTCCAGGCTTGGCAG	900
Db	1112	AAGAAGCCTGTTCAGATGCATCCCTCAGTTCGATCCGTTTCACCTTCCAGGCTTGGCAG	1171
Qy	901	CTGGCTACCAAGAGAAAGTTGCCATTTGCCAAGAAGTGAAGGATCGGGCGAAGAGGG	960
Db	1172	CTGGCTACCAAGAGAAAGTTGCCATTTGCCAAGAAGTGAAGGATCGGGCGAAGAGGG	1231
Qy	961	AGCACCGACACGACGATTCGAGTTCCGAGAAAGGAGCTTCGAGAAAACAAATAAGGGCAGT	1020
Db	1232	AGCACCGACACGACGATTCGAGTTCCGAGAAAGGAGCTTCGAGAAAACAAATAAGGGCAGT	1291
Qy	1021	TGTCGAGGCTTCTCGCAAGAGAAAGAGAGTGCAAAACATCCACGCGGAAGCGCCCTCC	1080
Db	1292	TGTCGAGGCTTCTCGCAAGAGAAAGAGAGTGCAAAACATCCACGCGGAAGCGCCCTCC	1351
Qy	1081	GTGACCGGGATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCCATCTACTCTCTCCGAGTGC	1140
Db	1352	GTGACCGGGATCCCAAGTTTACCCAGTAGCAAGAGAGAGCCCATCTACTCTCTCCGAGTGC	1411
Qy	1141	GGCAAGCTTTTCAGAACTACACACAGCTGGTCTTTGCACCTCCAGGGTCCACAAGAAGGAC	1200
Db	1412	GGCAAGCTTTTCAGAACTACACACAGCTGGTCTTTGCACCTCCAGGGTCCACAAGAAGGAC	1471
Qy	1201	CGGAGGGCCGGCGGGAGTCCGCCACATGTTCTGTGACGGGAGGACGCGGGGACGTGT	1260
Db	1472	CGGAGGGCCGGCGGGAGTCCGCCACATGTTCTGTGACGGGAGGACGCGGGGACGTGT	1531
Qy	1261	TCTCTGACCTCGCGGCCCTCTCGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGT	1320
Db	1532	TCTCTGACCTCGCGGCCCTCTCGATGAAAATGGAGCCGTGGATCGAGGGGAAGGTGGT	1591
Qy	1321	TCTGAAGACGGATCTGAGGATGGGCTTCCGAAGGAATCCATCTGGATAAAAATGATGAT	1380
Db	1592	TCTGAAGACGGATCTGAGGATGGGCTTCCGAAGGAATCCATCTGGATAAAAATGATGAT	1651
Qy	1381	GGAGGAAAAATAAACATCTTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTC	1440
Db	1652	GGAGGAAAAATAAACATCTTTACATCTTCAAGAGAGTGTAGTTATTGTGGAAAGTTTTC	1711
Qy	1441	CGTTCAAAATTAATACCTCAATATTTCACTCTCAGAACGCATACAGGTGAAAACCATACAAA	1500
Db	1712	CGTTCAAAATTAATACCTCAATATTTCACTCTCAGAACGCATACAGGTGAAAACCATACAAA	1771
Qy	1501	TGTGAATTTTGTGAATATGCTGCAGCCCGACAGATCATCTCTGAGGTATCATCTTGGAGAGA	1560
Db	1772	TGTGAATTTTGTGAATATGCTGCAGCCCGACAGATCATCTCTGAGGTATCATCTTGGAGAGA	1831
Qy	1561	CATCACAGGAAAAACAAACCGATGTTGCTCTGAGTCAAGACGATGTTAAAAATCAG	1620
Db	1832	CATCACAGGAAAAACAAACCGATGTTGCTCTGAGTCAAGACGATGTTAAAAATCAG	1891
Qy	1621	GACACTGAAGATGCACATTAATTAACCGCTGCACAGTGGCAAAACCAAAAAATTTTCAAAAGATTT	1680
Db	1892	GACACTGAAGATGCACATTAATTAACCGCTGCACAGTGGCAAAACCAAAAAATTTTCAAAAGATTT	1951
Qy	1681	TTTGATGTGCGAAAGATGTTTACAGGCGATGCCACTCTGCAAAAGAGCTTTAAGGAGATCGCT	1740
Db	1952	TTTGATGTGCGAAAGATGTTTACAGGCGATGCCACTCTGCAAAAGAGCTTTAAGGAGATCGCT	2011

Qy	1741	TCTGTTTTTTCAGAAATGTTCTGGGCAGCGTGTCTCTCACCGACGACACAAAGATACCTCAG	1800
Db	2012	TCTGTTTTTTCAGAAATGTTCTGGGCAGCGTGTCTCTCACCGACGACACAAAGATACCTCAG	2071
Qy	1801	GATTTCCATAAAAAATGCAGCTGATGACAGTGTCTGATAAAAGTGAATAAAAAACCTCACCCCT	1860
Db	2072	GATTTCCATAAAAAATGCAGCTGATGACAGTGTCTGATAAAAGTGAATAAAAAACCTCACCCCT	2131
Qy	1861	GCTTTACTCGTGCACCTGTTAAAAAAGAGATCAGCAGTTGTAAGAACTCAGGCAAAATACCTCATC	1920
Db	2132	GCTTTACTCGTGCACCTGTTAAAAAAGAGATCAGCAGTTGTAAGAACTCAGGCAAAATACCTCATC	2191
Qy	1921	TGTAGAACCAAGCGGAGATGTTACTCTCTCCGGATGGCAGTACACCCATACCTTTGAA	1980
Db	2192	TGTAGAACCAAGCGGAGATGTTACTCTCTCCGGATGGCAGTACACCCATACCTTTGAA	2251
Qy	1981	GTTAGCCCCCAAGAGAGCAAAACGGAGACGGCAGCTGACTGCAGATACAGGCCCAAGTGTG	2040
Db	2252	GTTAGCCCCCAAGAGAGCAAAACGGAGACGGCAGCTGACTGCAGATACAGGCCCAAGTGTG	2311
Qy	2041	GATTTGTCACAAAAACCTTTAAAAATTTATCCGTGGGGGCTCTTCACAATTTGCCCGCAAT	2100
Db	2312	GATTTGTCACAAAAACCTTTAAAAATTTATCCGTGGGGGCTCTTCACAATTTGCCCGCAAT	2371
Qy	2101	TCTTTGAGTAAAAAGTTTGATTCCAAGTATCACCTGTGCCATTTTGTGTAACCTTCAAGACATTT	2160
Db	2372	TCTTTGAGTAAAAAGTTTGATTCCAAGTATCACCTGTGCCATTTTGTGTAACCTTCAAGACATTT	2431
Qy	2161	TATCCAGAAGTTTTTAATGATGCAACGAGACTTGGAGCATATAATACAACTCTGACGTTTCAT	2220
Db	2432	TATCCAGAAGTTTTTAATGATGCAACGAGACTTGGAGCATATAATACAACTCTGACGTTTCAT	2491
Qy	2221	AAAAAATGTCGAAACCAAGTCTCTTCTTTAGAAATGCGAGCTACCCGGATGCCCGCAGCGTTG	2280
Db	2492	AAAAAATGTCGAAACCAAGTCTCTTCTTTAGAAATGCGAGCTACCCGGATGCCCGCAGCGTTG	2551
Qy	2281	CTGGAAAAAGATGTGCTCTCCCTCTCTPAGTTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTC	2340
Db	2552	CTGGAAAAAGATGTGCTCTCCCTCTCTPAGTTTTCTGTAAACCCAAAGCCCAAGTCTGCTTTC	2611
Qy	2341	CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC	2400
Db	2612	CCGGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGGAAGCAGAGCCCTCTCGGGCCAGGC	2671
Qy	2401	AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAAGTCC	2460
Db	2672	AAGGCCCTCTGACTTCAGGGATAGACTCTAGCACTTTAGCCCCCAAGTAACTGAAAGTCC	2731
Qy	2461	CACAGACCACAGCAGAAATGTGGGGGTCCAAAGGGCCGCCACCAAGGCAACAGCAATCTGAG	2520
Db	2732	CACAGACCACAGCAGAAATGTGGGGGTCCAAAGGGCCGCCACCAAGGCAACAGCAATCTGAG	2791
Qy	2521	ATGTTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAAGACAAAAGACCCGAGACAAAA	2580
Db	2792	ATGTTTTCTTAAACCAAGTGTTCCTCTGCAACCGGATAAGACAAAAGACCCGAGACAAAA	2851
Qy	2581	TTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCGGCGAGCAGTAGTAACTCAATGGT	2640
Db	2852	TTGAAACCTCTTCCAGTAGTCTCTTCTCAGCCACCTCGGCGAGCAGTAGTAACTCAATGGT	2911
Qy	2641	TCCATCGACTACCCCGCAAGAACGACAGGCCCTGGGCACCTCCGGGAAGAGACTATTTTC	2700
Db	2912	TCCATCGACTACCCCGCAAGAACGACAGGCCCTGGGCACCTCCGGGAAGAGACTATTTTC	2971
Qy	2701	TGTAATCGGAGTGCAGCAATACTGCAAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG	2760
Db	2972	TGTAATCGGAGTGCAGCAATACTGCAAGCAGAAATTTGGTGAGGCCCTTCCAAAAAGACTG	3031
Qy	2761	AAGTCCAGCGTGGTTGCCCTTGAGTTGACAGCCCGGGGCCAATTAACAGAAAGGCTAT	2820
Db	3032	AAGTCCAGCGTGGTTGCCCTTGAGTTGACAGCCCGGGGGCCAAATTAACAGAAAGGCTAT	3091
Qy	2821	GACCTTCCCAAGTAACTATATGTCAGAGGCATCACATCACTGTTTACCGCAGGACTGTGTG	2880

Db 3092 GACCTTCCCAAGTACCATATGTCAGAGGCATCACATCACTGTACCGCAGGACTGTGTG 3151
Qy 2881 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTGCAATCT 2940
Db 3152 TATCGTGCAGCGCTGCTCCAAAACCAAGGTTCTGAGCTCCAGCGAGGTGCAATCT 3211
Qy 2941 CCAATATGCTGACTGTTTCAGAAAGCCCTATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3000
Db 3212 CCAATATGCTGACTGTTTCAGAAAGCCCTATGTGTGCTCCGGGCCACTTTACACTTGTGTG 3271
Qy 3001 CTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAG----- 3034
Db 3272 CTGCTGTGTAGTCCAGCATCCAGCTCGAGTTAGAAAGGAAAGGCTGTGTGCATATCAA 3331
Qy 3035 ----- 3034
Db 3332 CACTTATCTAACAGCATGSCACAAAAGAGAAACTATGAGAAATTTATTGGGAATGCACAT 3391
Qy 3035 -----AAGGTCTTTGGTGG 3047
Db 3392 TATGACCAAAATGACAAAAAACTTGATTCTACTAATTAGGGGGAAAAAAGGTCITGGTGG 3451
Qy 3048 ATGTCAGTGTCTACTCCCATGAAATTTAAATTTTACTTTCATCTTTGAGAAAGCAATGCT 3107
Db 3452 ATGTCAGTGTCTACTCCCATGAAATTTAAATTTTACTTTCATCTTTGAGAGCGAATGCT 3511
Qy 3108 GAAAGCTACTGAAATTAAGCTGTGATGTACTGTACATAAAACATATGAGGAATCTGCAAG 3167
Db 3512 GAAAGCTACTGAAATTAAGCTGTGATGTACTGTACATAAAACATATGAGGAATCTGCAAG 3571
Qy 3168 GAACACTACAGTTGTGTAA 3186
Db 3572 GAACACTACAGTTGTGTAA 3590

RESULT 4

US-10-087-192-1708
; Sequence 1708, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIORITY FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1708
; LENGTH: 36022
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-087-192-1708

Query Match 49.0%; Score 1561.2; DB 13; Length 36022;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1566; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 1541
Db 15816 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCCAAGAGACATCTCT 15875
Qy 1542 GAGGTATCACTTGGAGAGACATCACAGGAAAAACAAACCGATGTGTGCTGTAAGTCAA 1601
Db 15876 GAGGTATCACTTGGAGAGACATCACAGGAAAAACAAACCGATGTGTGCTGTAAGTCAA 15935
Qy 1602 GAACGATGTTAAAAATCAGGACATGAAGATGCATATTAAACCGCTGACAGTGGCGAAAC 1661

Db 15936 GAACGATGTTAAAAATCAGGACATGAAGATGCATATTAAACCGCTGACAGTGGCGAAAC 15995
Qy 1662 CAAAAATTTGAAAGATTTTTTGATGGTCCAAAGATGTTTACAGCAGTCCACCTGCAAA 1721
Db 15996 CAAAAATTTGAAAGATTTTTTGATGGTCCAAAGATGTTTACAGCAGTCCACCTGCAAA 16055
Qy 1722 GCAGCTTAAGGAGATGCTTCTGTTTTTTCAGAAATGTTTCTGGCGAGGCTGCTCTCACC 1781
Db 16056 GCAGCTTAAGGAGATGCTTCTGTTTTTTCAGAAATGTTTCTGGCGAGGCTGCTCTCACC 16115
Qy 1782 AGCACACAAAGATACTCAGGATTTTCCATAAAAAATCAGCTGATGACAGTGTGATAAAGT 1841
Db 16116 AGCACACAAAGATACTCAGGATTTTCCATAAAAAATCAGCTGATGACAGTGTGATAAAGT 16175
Qy 1842 GAATAAAAACCTTACCTCTGCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 1901
Db 16176 GAATAAAAACCTTACCTCTGCTTACCTGGACCTGTTTAAAAAGAGATCAGCAGTTGAAAC 16235
Qy 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTTACTCTCTCCGATGGGAG 1961
Db 16236 TCAGGCAATTAACCTCATCTGTAGAACCAAGCGGATGTTTACTCTCTCCGATGGGAG 16295
Qy 1962 TACCACCCATTAACCTTGAAGTTTAGCCCCAAAGAGAGAAAGCAACGAGACCGCAGCTGACTG 2021
Db 16296 TACCACCCATTAACCTTGAAGTTTAGCCCCAAAGAGAGAAAGCAACGAGACCGCAGCTGACTG 16355
Qy 2022 CAGATACAGGCCAAGTGTGGATGTGCAGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 2081
Db 16356 CAGATACAGGCCAAGTGTGGATGTGCAGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 16415
Qy 2082 TCACAAATTTGCCCGGCAATTTCTTTGAGTAAAGATTTGATCCAGTATCATCCTGTCCATT 2141
Db 16416 TCACAAATTTGCCCGGCAATTTCTTTGAGTAAAGATTTGATCCAGTATCATCCTGTCCATT 16475
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTATGATGACACAGAGACTGGAGCATAA 2201
Db 16476 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTATGATGACACAGAGACTGGAGCATAA 16535
Qy 2202 ATACAATCTTGACGTTTCATAAAAACTGTCGAAAAAGTCTTGTGTTAGAAAGTTCGACGTAC 2261
Db 16536 ATACAATCTTGACGTTTCATAAAAACTGTCGAAAAAGTCTTGTGTTAGAAAGTTCGACGTAC 16595
Qy 2262 CGATGCCCCGCGAGGTTGCTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACC 2321
Db 16596 CGATGCCCCGCGAGGTTGCTGGGAAAGATGTGCTCCCTCTCTCTAGTTTCTGTAAACC 16655
Qy 2322 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAAGCA 2381
Db 16656 CAAGCCCAAGTCTGCTTTCCCGCGCAGTCCAAATCCCTGCCATCTGCGAAGGGAAGCA 16715
Qy 2382 GAGCCCTCTCTGGGCCAGGCAAGGCCCTCTGACTTTCAGGATAGACTCTAGCATTATTAGC 2441
Db 16716 GAGCCCTCTCTGGGCCAGGCAAGGCCCTCTGACTTTCAGGATAGACTCTAGCATTATTAGC 16775
Qy 2442 CCCAAGTAACTGAAATCCCAAGACACACAGAGATGTGGGGTCCAAAGGGGCGCCAC 2501
Db 16776 CCCAAGTAACTGAAATCCCAAGACACACAGAGATGTGGGGTCCAAAGGGGCGCCAC 16835
Qy 2502 CAGCAACAGCAATCTGAGATGTTTCTTAAACCCAGTGTTCCTCCCTGCCACCGGATAAGAC 2561
Db 16836 CAGCAACAGCAATCTGAGATGTTTCTTAAACCCAGTGTTCCTCCCTGCCACCGGATAAGAC 16895
Qy 2562 AAAAAGACCCGAGACAAAAATTTGAAACCTTCTCCAGTAGCTCTTCTCAGCCCCACCCTCGG 2621
Db 16896 AAAAAGACCCGAGACAAAAATTTGAAACCTTCTCCAGTAGCTCTTCTCAGCCCCACCCTCGG 16955
Qy 2622 CAGCAGTAAATCAATGGTTCCATCGACTACCCCGCCAAAGAAACGACACCCCTGGGAC 2681
Db 16956 CAGCAGTAAATCAATGGTTCCATCGACTACCCCGCCAAAGAAACGACACCCCTGGGAC 17015
Qy 2682 TCCGGAGAGACTATTCTGTATTCGAGTGCAGCAATACTCGACAGAAATTTGGTGA 2741

Db 17016 TCCGGGAAGAGACTATTCTCTGTAATCGGAGTGCACGCAATACTGACGAGAAATTTGGTGA 17075
Qy 2742 GCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTTGACCAAGCCGGGGC 2801
Db 17076 GCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGACGTTGACCAAGCCGGGGC 17135
Qy 2802 CAATTACAGAAGAGCTATGACCTTCCCAAGTACCATATGCTCAGAGCATCATCACT 2861
Db 17136 CAATTACAGAAGAGCTATGACCTTCCCAAGTACCATATGCTCAGAGCATCATCACT 17195
Qy 2862 GTTACCGCAGACTGTGTATTCGTCGACGGCGTGCCTCCAAACCAAGTTCCTGAG 2921
Db 17196 GTTACCGCAGACTGTGTATTCGTCGACGGCGTGCCTCCAAACCAAGTTCCTGAG 17255
Qy 2922 CTCAGCGAGGTGCATTCTCCAAATGTGCTGACTGTTCCAGAAGCCCTATGCTGCTCCGG 2981
Db 17256 CTCAGCGAGGTGCATTCTCCAAATGTGCTGACTGTTCCAGAAGCCCTATGCTGCTCCGG 17315
Qy 2982 GCCACTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3041
Db 17316 GCCACTTTACACTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 17375
Qy 3042 TGGTGGATGTCAGT 3055
Db 17376 TGCATGAGGGCGT 17389

RESULT 5

US-08-731-499-9
; Sequence 9, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION:
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin
; APPLICANT: HWANG, Soo-In
; APPLICANT: GODFREY, Tony
; APPLICANT: KOWBEL, David
; APPLICANT: ROMMENS, Johanna
; TITLE OF INVENTION: GENES FROM THE 20q13 AMPLICON AND THEIR
; TITLE OF INVENTION: USES
; NUMBER OF SEQUENCES: 44
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,499
; FILING DATE: 16-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/680,395
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Hunter, Tom
; REGISTRATION NUMBER: 38,498
; REFERENCE/DOCKET NUMBER: 23070-068910
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10365 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: 1..10365
; LOCATION: 1..10365
; OTHER INFORMATION: /note= "Genomic Sequence Encoding
; OTHER INFORMATION: ZABC1"
; US-08-731-499-9

Query Match 49.0%; Score 1559.6; DB 8; Length 10365;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1565; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 1482 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT 1541
Db 8286 AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGCAGCCAGAGACATCTCT 8345
Qy 1542 GAGGTATCACTTGCAGAGACATCAAGAGAAAAACAACCGATGTTGCTGCTGAAGTCAA 1601
Db 8346 GAGGTATCACTTGCAGAGACATCAAGAGAAAAACAACCGATGTTGCTGCTGAAGTCAA 8405
Qy 1602 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAAC 1661
Db 8406 GAACGATGGTAAAAATCAGGACACTGAAGATGCACTATTAAACCGCTGACAGTGGCGAAAC 8465
Qy 1662 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTACAGGAGTCCACCTGCAAA 1721
Db 8466 CAAAAATTTGAAAAAGATTTTGTGATGTGCAAAAGATGTTACAGGAGTCCACCTGCAAA 8525
Qy 1722 GCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 1781
Db 8526 GCAGCTTAAGGAGATGCTCTGTTTTTTCAGAAATGTTCTGGGAGCGCTGCTCTCACC 8585
Qy 1782 AGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 1841
Db 8586 AGCACACAAAGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTGTGATAAAGT 8645
Qy 1842 GAATAAAAAACCTACCCCTGCTTACCTGGAGCCTGTTAAAAAAGAGATCAGCAGTTGAAAC 1901
Db 8646 GAATAAAAAACCTACCCCTGCTTACCTGGAGCCTGTTAAAAAAGAGATCAGCAGTTGAAAC 8705
Qy 1902 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGATGGCAG 1961
Db 8706 TCAGGCAATTAACCTCATCTGTAGAACCAAGGCGGATGTTACTCTCTCCGATGGCAG 8765
Qy 1962 TACCACCCATAACTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 2021
Db 8766 TACCACCCATTAACCTTGAAGTTAGCCCCCAAGAGAAAGCAAAACGGAGACCGCAGCTGACTG 8825
Qy 2022 CAGATACAGGCCAAGTGTGGATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 2081
Db 8826 CAGATACAGGCCAAGTGTGGATTGTACGAAAAACCTTTTAAATTTATCCGTGGGGGCTCT 8885
Qy 2082 TCACANTTCCCGGCAATTTCTTGTAGTAAAGTTGATTCCAAGTATCACTGTCCATT 2141
Db 8886 TCACAAATTCGCGGCAATTTCTTGTAGTAAAGTTGATTCCAAGTATCACTGTCCATT 8945
Qy 2142 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCCAGAGACTGGAGCATAA 2201
Db 8946 TTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGACCCAGAGACTGGAGCATAA 9005
Qy 2202 ATACAAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCCCTGCTTAGAAGTTCGACGAC 2261
Db 9006 ATACAAATCCTGACGTTTCATAAAAACTGTGCAAAACAAAGTCCCTTGTAGAAAGTTCGACGAC 9065
Qy 2262 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCTCCCTCTCTAGTCTCTGTAACC 2321
Db 9066 CGGATGCCCGCAGGTTGCTGGGAAAGATGTGCTCCCTCCCTAGTCTCTGTAACC 9125
Qy 2322 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAAGCA 2381
Db 9126 CAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCAATCTGCGAAGGGGAAGCA 9185
Qy 2382 GAGCCCTCCTGGGCGCAGGCAAGGGCCCTCTGACTTTCAGGGATAGACTCTAGCACTTTAGC 2441

Db 9186 GAGCCCTCTCGGCGCAGGCAAGCCCTCTGACTTCAGGATAGACTCTAGCACTTTAGC 9245
Qy 2442 CCCAAGTAACCTGAAGTCCACAGACCCAGCAGCAAGTAATGTGGGGTCCAAAGGGCGCCGAC 2501
Db 9246 CCCAAGTAACCTGAAGTCCACAGACCCAGCAGCAAGTAATGTGGGGTCCAAAGGGCGCCGAC 9305
Qy 2502 CAGGCAACAGCAATCTGAGATGTTTCTTAACACCAAGTGTTCCTCCCTGACCGGATAGAC 2561
Db 9306 CAGGCAACAGCAATCTGAGATGTTTCTTAACACCAAGTGTTCCTCCCTGACCGGATAGAC 9365
Qy 2562 AAAAGAGCCCGACAGACAAATTTGAACCTCTTCCAGTAGTCTCTTCCAGCCACCCCTCGG 2621
Db 9366 AAAAGAGCCCGACAGACAAATTTGAACCTCTTCCAGTAGTCTCTTCCAGCCACCCCTCGG 9425
Qy 2622 CAGCAGTAACATCAATGTGTTCCATCGACTACCCGCCAAGACGACAGCCGCTGGGCACC 2681
Db 9426 CAGCAGTAACATCAATGTGTTCCATCGACTACCCGCCAAGACGACAGCCGCTGGGCACC 9485
Qy 2682 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGACGAGTAATTTGGTGA 2741
Db 9486 TCCGGGAAGAGACTATTTCTGTAATCGGAGTGCAGCAATACTGACGAGTAATTTGGTGA 9545
Qy 2742 GCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTCGCCCTTGACGTTGACCGCCCGGGC 2801
Db 9546 GCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTTCGCCCTTGACGTTGACCGCCCGGGC 9605
Qy 2802 CAATTACAGAGAGGCTATGACTTCCCAAGTACCATATGTTGACAGGATCATCACT 2861
Db 9606 CAATTACAGAGAGGCTATGACTTCCCAAGTACCATATGTTGACAGGATCATCACT 9665
Qy 2862 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCTGCTCCCAACCAAGTTCCTGAG 2921
Db 9666 GTTACCGCAGGACTGTGTATTCCTCGCAGCGCTGCTCCCAACCAAGTTCCTGAG 9725
Qy 2922 CTCACGAGGCTCGATTCTCAAATGTCTGACTGTTCAGAGCCCTATGTTGGCTCCGG 2981
Db 9726 CTCACGAGGCTCGATTCTCAAATGTCTGACTGTTCAGAGCCCTATGTTGGCTCCGG 9785
Qy 2982 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTGTAGAGGTCT 3041
Db 9786 GCCACTTTACACTTGTGTCCTGCTGGTAGTCCAGCATCCAGCTCGAGTGTAGAGGTAT 9845
Qy 3042 TGGTGGATGTCAGT 3055
Db 9846 TGCATGAGGGCGGT 9859

RESULT 6

US-10-087-192-1706
; Sequence 1706, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIORITY FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIORITY FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIORITY FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1706
; LENGTH: 3016
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1706

Query Match

47.4%; Score 1511.2; DB 13; Length 3016;

Best Local Similarity 71.5%; Pred. No. 0;
Matches 2194; Conservative 6; Mismatches 759; Indels 111; Gaps 12;
Qy 25 ATGCAACTCAATCCCTCTTAATGATGATGAGGCGCAGAGTGAATGGCAGCTCTCTT 84
Db 1 ATGCGGACCCAGTCCCTCTCGTGTACATGAGCGGCGCGAAGTCTCTCAGCAGCTCTCTA 60
Qy 85 GGCAGTCCGATGAGATGGAGATGCCCTTGTCAATGAAGGAGACCGCTGTGTTGCATTC 144
Db 61 GGCCTCCAGATGAGAGGTGGATGTCTGTGCCCATAAAAGGGCGCGTGGCAGTCCCTTC 120
Qy 145 CGAGCTACACAAGAAAAAATGTTCATCCAAATCGAGGGGTATATCCCTTGGATGGCATG 204
Db 121 CGAGCTGTCTAGGAGAGAGATGGCCGTGGCAGAGGGCCCATGCCCTCTGATGGATG 180
Qy 205 TTCTGCGCCAGACCTTCCACACATTTAGAAGACCTTAATAAACAATGTCTTAATGAACAC 264
Db 181 TTCTGCGCCAGGCTCTTCTCTCAGCGGAGGATCTCACTCAGCAGCTGTCTGTCGACAC 240
Qy 265 CGGCTACCTCTGTGAACACGAGCTTCTTGGGTTGAAGCAGATATCTCAGTCCGCTT 324
Db 241 CGGCCACCTCTCGAGCCAGCTGTCTGCGTGTGGAGCGCGAGTACCTTAAGTCCCTT 300
Qy 325 GATAAAGTCAAGTGCAGAACAGAACCTCCCAAGGAAAAAATGCAAGAA --- AAATGAA 381
Db 301 GATAAAGCTCTGGAGCCACAGAGCCACATTTGGAGAGAGTGGCGAAGACCCCGAGGAG 360
Qy 382 TTTAGCTGTGAGGTATGTGGGCAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCACATG 441
Db 361 TTGAGCTGTGATGTGTGGGCAGACATTTCCAGTGGCTTTTGTGATGTTGAGAGCCACATG 420
Qy 442 AGAACACAAAGATTTTTCATCTTAACCTTACCGGTGTAAATGTCGGGAAGAGMTTSSRSAG 501
Db 421 AAGAAGCATAAAGGACTCTCTTCAAGTATGGGTGACAGCATGTGCGGAGGAGATTTCAAGAG 480
Qy 502 CTTGGTGTCTTAAAAATCAGATCGGACACATATGCAAAATCGGGGCGCAGAGAGCAA 561
Db 481 CCGTGTGTTCTGAGAACCAATCGGACACACATATGGCAAGTCTGGCACACAGGAGCAG 540
Qy 562 CTGACGCAAGGCTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGACCGG 621
Db 541 CTTCAGCAAGGATGAGG --- AGTCCAGTCAACATCAATGAAGTGTCCAGCGCACGCC 597
Qy 622 GCGAGAGCATCTCTCTCTTCAAAATCTGATGTTGTTGGCTTCTTATTTCCAAAT 681
Db 598 CTTGGAGCATCTCCACGCTTACAGATCTGATGTTGTTGGCTTCTCTTCTTCCAAAT 657
Qy 682 AAAGAAAGTCTAAATTTGAGCACCGAAGGTGCACACCAAAATAAAGTCTTTCGTTACCGC 741
Db 658 AAGCAGAGCCCTCTTTGAGCAGCAGCAAGGTTTCAAGCAAAAGAACTGTCCCGAGTCCGAG 717
Qy 742 AGCGCGCAGACAGACTCTTCCACAAGGAGGAATGCGCTCTCGAGGAGGAGTCTTCTGCGAG 801
Db 718 AACGTTGCTGATGATGACCCAGAGAGAAACCCAGCTCCCGAGGGAAGAGTGTGTCGAG 777
Qy 802 TTGTTCAACTTTGAGACCAAAATCTCACCTGAAACGGGGAAGAGCTGTGATGATC 861
Db 778 TTTTGAACCTTTGAGACCCAGATCAACTCAGGTAGTACAGTGAAGCCCATGACCTGCATA 837
Qy 862 CCTCAGCTCGATCCGTTTCAACCACTTCCAGCTTGGCAGCTGGCTTACCAAGGAAAGT 921
Db 838 CCTCAGCTTGAACCCGTTTCAACCACTTCCAGGATGAGGAGTGGCTTACCAAGGAAAGT 897
Qy 922 GCCATTTTGC --- AAGAGTGAAGGAATCGGGGCAAGAGGAGCACCAGCAACACGAGT 978
Db 898 GCGGTTGCCAGGAAGAGGTGAAGAGTCAAGGCAAGAGGAGGAGCAGACATCAACGAC 957
Qy 979 TCGAGTTCCGAGAGGAGCTTGGAG --- 1003
Db 958 TCATGCTCAGAGAAAGAGGAACTAGGGGAAATATGGGTTGGGGTAAGCGGAGAGGTCT 1017
Qy 1004 -----AAACAATAAGGAGGAGTGTGCGGCTCTCGCAGGAAAGAGAGTGC 1053

Db 1018 GGAAGTCCAAAAAAGTAAAGCAGTTGTCCAGGTCTCTCCCAAGACAAAGAGAGCCT 1077
Qy 1054 AAACACTCCACGGCGAGCGCCCTCCGTGACCGGATCCCAAGTTACCAGTAGCAAG 1113
Db 1078 AGACATGCTAATAGTGAAGTCCCTTCTGGGATAGTAGCCCAAGTTGTCCAGTAGCAAG 1137
Qy 1114 GAGAAAGCCACTCACTGCTCCGAGTGGCGGCAAGCTTTTCAGAACTACCAACAGCTGGTC 1173
Db 1138 GAGAAAGCCACGCACTGTTCTGAGTGCAGCAAGGCTTCAGGACATACCAACAGCTGGTC 1197
Qy 1174 TTGCATCCAGGGTCCACAAAGAGACCGGAGGCGCGGAGTGGCCGACCATGTCT 1233
Db 1198 CTGCACCTGAGGGTGCACAGGAAGCAGAGGAGTCTGATGCGCTGTCAACCAACATGGCT 1257
Qy 1234 GTGACGGGAGGACCGGGAGCTGTCTCTGACCTCGCGCGCTCTGATGAAAT 1293
Db 1258 GTGATGCAAGGACGCTGGAGCCTGTCTCCAGACCTCAGCACTCTGGAAGACAGT 1317
Qy 1294 GGAGCGGTGGATCGAGGGGAAGTGTCTGAAAGCGGATCTGAGGATGGGCTTTCGGA 1353
Db 1318 GGGCCGGGACCGA--GAAGGGGCTCTGAAGCGGCTCTGAGGATGGACTCCCTGAC 1374
Qy 1354 GGAATCCATCTGGATATAAATGATGAGAGGAAATAAACAATCTTACATCTTCAGA 1413
Db 1375 GGGCTCCATTTGGATATAAATGATGAGGAAAGAGCGCCCTCCCTCCCTCGAGA 1434
Qy 1414 GAGTGTAGTTATTGTGGAAGCTTTTCGTTCAATTTATACCTCAATTTCTCTCAGA 1473
Db 1435 GAGTGTAGTTACTGTGGCAAGTTTTCGTTCAAACTATTAACCTCAATTTCTCTCAGA 1494
Qy 1474 ACGCATACAGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGACGCCAGAG 1533
Db 1495 ACGCATACAGGTGAAACCAATACAAATGTGAATTTTGTGAATATGCTGACGCCAGAG 1554
Qy 1534 ACATCTCTGAGGTATCACTTGGAGAGACATCACAAAGGAAACAAAC---CGATGTGCT 1590
Db 1555 ACATCTCTGAGGTACCACTTGGAGAGACATCACAAAGAGCAAGCAGCGCGGTGGATGCTGCC 1614
Qy 1591 GCTGAAGTCAGAACGATGTTAAATTCAGACACTGAAGTGCACATTTAA---CCGCT 1647
Db 1615 GCTGAGTCCAAAGTGAAGCGCGAGCCAGGCGCGAGGATGGCTACTAACCGCTGCT 1674
Qy 1648 GACAGTGGCAAAACCAAAATTTGAAAGATTTTGTGATGTTGCGCAAGATGTTACAGGC 1707
Db 1675 GACAGTGGCAAGCAAAATTTAAAGAGATTTCTGTGATGTCGCAAGATGTTAAAGGA 1734
Qy 1708 AGTCACCTGCAAGCAGCTTAAGGAGATGCTCTGTTTTTCAGATGTTCTGGGCAGC 1767
Db 1735 AGCCACCTGCCAAGCAGCTTAAGGAGATGCTCTGTCTTCCAGAGTGTTC----- 1786
Qy 1768 GCTGCTCTCACAGCACACAAAGATACTCAGGATTTCCATTAATAATGACGCTGATGAC 1827
Db 1787 ----TCTCACCAGCACACAGCAACGATACTCAGGATTTCCATAACATGACGCTGAT--- 1839
Qy 1828 AGTGTGATAAGTGAATAAAACCCCTACCCCTGCTTACCTGGACCTGTATAAAAGAGA 1887
Db 1840 AGTGTGATAAGCGAGGAAGAGCCCTGCCCCCTACTTATCTGGACATCCAGGAAGA-- 1897
Qy 1888 TCAGAGTTGAACTCAGGCAAAATTAACCTCATCTGTAGAAACAAGCGGATGTTACTCCT 1947
Db 1898 -AAGCAGGGAGCCTCAGGCCAGCAGCCCTGTCTGCAGACTAGAGGGGTTGGGTCTCTTA 1956
Qy 1948 CTTCCGGATGGCAGTACCACCATACCTTGAAGTTAGCCCCCAAGAGAGCAACGGAG 2007
Db 1957 GCACGGGAAGCTGG-----CCATAGGGGAGAAAGTGGAT 1989
Qy 2008 ACCGAGCTGACTGCAGATACAGGCCAAGTGTGGATTTGTACGAAAAACCTTTTAAATTTA 2067
Db 1990 CAGGATGCTACTACAGACATTAAGCCCGTGTGACTGCCAGGACAGGCCCTTGAATCTA 2049
Qy 2068 TCCGTGGGGGCTTTTCAAAATTTGCCGGCAATTTCTTTGAGTAAAGTTTGAATCCAAGT 2127
Db 2050 TCCCTTGGGCGGCTCCAGGCCTGTCTCTGCAATCTCTTTGAGCAAGTGTCTGATCCCGAGC 2109

Qy 2128 ATCACTGTCTCAATTTTGTACCTTCAAGACATTTTATCCAGAAGTTTAAATGATGCACAG 2187
Db 2110 ATTGCTGCCCCCTTTTGTACTTTCAAGACCTTTTATCCGGAAGTCTTATGATGCACAG 2169
Qy 2188 AGACTGGAGCATATAAATAAATCTCTGAGCTTTATATAAACTGTTCGAAAACAAGTCTTGCTT 2247
Db 2170 AGACTTGAGCACAGGTACAAACCTTGACCCGACAAAGAAACGCGCAGCAGCAAGTCTGTGCTG 2229
Qy 2248 AGAAGTCAGTACCGGATGCCCGCAGCGTTGTCTGGGAAAGATGTGCCCTCCCTCTCT 2307
Db 2230 AGGAACAGGGGTACCGGGTCCCTCCGGCTTTGCTGGGAAAGATGTCCCTCCCTTGTCT 2289
Qy 2308 AGTTTCTGTAAACCAAGCCCAAGCTCTGCTTTTCCCGCGCAGTCCAAATCCCTGCCATCT 2367
Db 2290 GGCTTCGACAAAGCCCAAGCCNAGACTGCCCTTCAACCACTCGAAGTCCCTGCACCTCA 2349
Qy 2368 GCGAAGGGGAAGCAGAGCCCTCTCTGGCCAGGCAAGGCCCTCTGACTTCAGGATAGAC 2427
Db 2350 GAGAAGGCTCGGCAGGGGCTCGGGGCCAAGCAAGCAAGCACCACAGACTTCAGGACCCAGAC 2409
Qy 2428 TCTAGCACTTTAGCCCCCAAGTTAAGCTGAAGTCCACAGACACACAGCAAGATGTGGGGTC 2487
Db 2410 AACAGCACTTTAGCCCCCAAGTTAAGTCAACAGGTCAACAAACCAATGCTGGGGGC 2469
Qy 2488 CAAGGGCGCGCCACCGCAACAGCAATCTGAGATGTTTCTTAAACCCAGTGTTCCTCCCT 2547
Db 2470 ACCAG--TGGCCACACGCGCAGCAGTCAAGTTGTTTCCCAAGGTGGCGTCCCTGCT 2526
Qy 2548 GCACCGGATTAAGCAAAAGAACCCGAGACAAATTTGAAAACCTCTTCCAGTAGTCTCTTCT 2607
Db 2527 GCTATGATTAAGTGAAGAGACTGAGCCAAAACCTGAAGTCCCTTACCAGCCTCCCGCTCT 2586
Qy 2608 CAGCCCACTCTCGCAGCAGTAAACATCAATGGTTTCCATCGACTACCCGCGCAAGACGAC 2667
Db 2587 CAGTCCCCCTCAGCAGTAATAAGCAACGGTTCTGTGTAGTATCCCGTGAAGTTGAC 2646
Qy 2668 AGCCCGTGGGCACCTCGGGGAAGAGACTATTCTGTATTCGGAGTGCAGCAATACTGCA 2727
Db 2647 GGCCTATGGGCACAGCAAGGGAGAGACTACTACTGCCATCGGAATCTTGGCAGTGCAGCA 2706
Qy 2728 GCAGAAATTTGTGAGCCCCCTTCCAAAAGACTGAAGTCCAGCGTGGTGGCCCTTGAAGT 2787
Db 2707 GCAGAGTACAGTGAAGCAGACATCCCAAAAGACTCAAGTCCAGTGGCGGTGCTCCCTGGACACA 2766
Qy 2788 GACCAGCCGGGCGCAATTACAGAGAGGCTATGACCTTCCCAAGTACCATATGTCAGA 2847
Db 2767 GAGCATGACGGGACCAATGGCAGAAAGGGGCTTTGAGCTCCCCCAAGTACCATGTGTGTCAGG 2826
Qy 2848 GGCAATCACATCACTGTTTACCGCAGGACTGTGTATTCGTCGCGAGGCGCTGCCTCCCAAA 2907
Db 2827 AGCATCACTCTTGTCTACCAACAGAGTGTGTGGCCACCGCTGTGCTGCCCCACAAA 2886
Qy 2908 CCAAGGTTCTGAGCTCCAGGAGTGCATTTCTCAATGTGCTGACTGTTCAGAGGCC 2967
Db 2887 GCGCGTTTCTCTGAGCCCTCGGGAGGTGAGTCAACCCAGTGTGTGGCTGTGCGAAGGCC 2946
Qy 2968 TATGTTGCTCCGGGCGCACTTTTACACTGTGTGCTGCTGTGTCAGACTCCAGCTCG 3027
Db 2947 TACAGTGCCTCTGAGACCCCTGTATACCTGTGTGACCCCGTGGGACACGCGAGGAGCGACCA 3006
Qy 3028 ACGTTAGAAG 3037
Db 3007 GCGCTTGAAG 3016

RESULT 7
US-08-731-499-3
; Sequence 3, Application US/08731499
; Publication No. US20030148270A1
; GENERAL INFORMATION: Joe W.
; APPLICANT: GRAY, Joe W.
; APPLICANT: COLLINS, Colin


```

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1705
; LENGTH: 26345
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-087-192-1705

```

Query Match 23.0%; Score 733; DB 13; Length 26345;
Best Local Similarity 70.2%; Pred. No. 7.5e-206;
Matches 1106; Conservative 0; Mismatches 415; Indels 54; Gaps 7;

Qy	1482	AGGTGAAAAACCATACAAATGTGAATTTTGTGAATATGCTGAGCCGACAGACATCTCT	1541
Db	14832	AGGTGAAAAACCATACAAATGTGAATTTTGTGAATGCGCAGCCAGACATCTCT	14891
Qy	1542	GAGGTATCATCTCGAGAGACATCACAAAGGAAAAAACAAC--CGATGTTGCTGCTGAAGT	1598
Db	14892	GAGGTATCATCTCGAGAGACATCACAAAGGAAAAAACAAC--CGATGTTGCTGCTGAAGT	14951
Qy	1599	CAAGAACGATGTTGAAAAATCAGGACACTGAAGATGCATTTATA---CGCTGACAGTGC	1655
Db	14952	CAAAAGTGAAGCGCGGACGAGCCGAGGATGCCTACTTAAACGGCTGCTGACAGTGC	15011
Qy	1656	GCAAAACCAAAATTTGAAAAGATTTTGTATGTTGCCAAAGATTTTACAGGAGTCCACC	1715
Db	15012	GCAGACCAAAATTTTAAAGAGATTTCTGATGTTGCCAAAGATTTTAAAGGAAGCCACC	15071
Qy	1716	TGCAAAAGACGCTTAAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCAGCGCTGCT	1775
Db	15072	TGCCAAGCAGCTTAAAGGAGATGCTTCTGTTTTCAGAAATGTTCTGGGCAGCGCTGCT	15119
Qy	1776	CTCACCAGCACAAAGAGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTCTGA	1835
Db	15120	ACCAGCACACAGCAACGATACTCAGGATTTCCATAAAAAATGCAGCTGATGACAGTCTGA	15176
Qy	1836	TAAAGTGAATAAAAACCTTACCTCCCTGTTACCTGGACCTGTTTAAAAAGAGATCAGCAGT	1895
Db	15177	GAAAGCGAGGAAGAGCCCTGCCCCCTACTTATCTGGACATGACAGAAAGAAAA---GCAGG	15233
Qy	1896	TGAAACTCAGGCAAAATAACCTCATCTGTAGAACCAAGCGGATGTTACTCTCTCCCGGA	1955
Db	15234	GGAGCCTCAGGCCAGACGCCCTCTCTGCAGCTAGAGGGGGTTGGTCTCTTAGCACGGGA	15293
Qy	1956	TGGCAGTACACCCCATAAACCTTGAAGTTAGCCCCAAAGAGAGACAAACGAGACCGCAGC	2015
Db	15294	AGCTGG-----CCATAGGAGAGATGATGATCAGGATGC	15326
Qy	2016	TGACTGCAGATACAGGCCAAGTGTGATTTGTACAGAAAAACCTTTAAATTTATCGTGGG	2075
Db	15327	TGACTACAGACATAAGACCCGGTCTGACTGCCAGGACAGCCCTTTGAAATCTATCCCTTGG	15386
Qy	2076	GGCTCTTCAAAATTCGCCCGCAATTTCTTTTGAGTAAAGTTGATTTCCAAAGTATCACCTG	2135
Db	15387	GCGCTCCAGCCTGTCTCTGCAATCTCTTTTGACAAAGTGTCTGATCCCGCAGATTCGCTG	15446
Qy	2136	TCCATTTTGTACCTTCAAGACATTTTATCCAGAAAGTTTAAATGATGCACACAGAGACTGA	2195
Db	15447	CCCTTTTGTACTTCAAGACCTTTTATCCGGAAGTCTTATGATGACACACAGAGACTGA	15506
Qy	2196	GCATAAATACAATCTGACGTTTCATAAAACTGTGAAAACAGTCTCTGTTAGAGTGG	2255
Db	15507	GCACAGGTACAAACCTGACCCGACCAAGACGCGCAGCAAGTCTGTGCTGAGGAACAG	15566
Qy	2256	ACGTACCGGATGCCCGCAGCGTTGCTGGGAAAGATGTCCTCCCTCTCTAGTTTCTG	2315
Db	15567	GCGTACCGGGTGCCCTCCGGCTTTGCTGGGAAAGATGTCCTCCCTCTCTGTTGGCTGCA	15626
Qy	2316	TAAACCAAGCCCAAGTCTGCTTTCCCGGCGCAGTCCAAATCCCTGCCACTCTGCGAAGG	2375

MEDIUM TYPE: Floppy Disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/040,739
FILING DATE: 07-Jan-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/036,520
FILING DATE: 03-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 520:
SEQUENCE CHARACTERISTICS:
LENGTH: 469 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 520:
US-10-040-739-520

Query Match 13.9%; Score 444.4; DB 13; Length 469;
Best Local Similarity 99.8%; Pred. No. 5.4e-121;
Matches 445; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1032 CTCGCAAGAGAGAGAGTGCACACCTCCACGCGGAGCGCCCTCCGTGGAGCGGA 1091
Db 19 CTAGCAAGAGAGAGAGTGCACACCTCCACGCGGAGCGCCCTCCGTGGAGCGGA 78
Qy 1092 TCCCAAGTTACCCAGTAGCAAGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCTTT 1151
Db 79 TCCCAAGTTACCCAGTAGCAAGAGAGCCACTCACTGCTCCGAGTGGCGCAAGCTTT 138
Qy 1152 CAGAACCTTACCAACAGTGGTCTTGCACTCCAGGGTCCACAGAGGACCGGAGCGCG 1211
Db 139 CAGAACCTTACCAACAGTGGTCTTGCACTCCAGGGTCCACAGAGGACCGGAGCGCG 198
Qy 1212 CGCGGAGTGGCCACCATGCTGTGGAGCGGAGGAGCGCGGAGCGTCTCTCTGACCT 1271
Db 199 CGCGGAGTGGCCACCATGCTGTGGAGCGGAGGAGCGCGGAGCGTCTCTCTGACCT 258
Qy 1272 CGCGCCCTCTTGATGAAATGGAGCGGAGGAGCGGAGGAGCGTCTCTGAGACGG 1331
Db 259 CGCGCCCTCTTGATGAAATGGAGCGGAGGAGCGGAGGAGCGTCTCTGAGACGG 318
Qy 1332 ATCTGAGATGGGCTTCCGAGGAGTCCATCTGGATGAAATGATGATGAGGAGGAGGAG 1391
Db 319 ATCTGAGATGGGCTTCCGAGGAGTCCATCTGGATGAAATGATGATGAGGAGGAGGAG 378
Qy 1392 AARACATCTACATCTTCAAGAGAGTGTAGTTATGCGAAAGTTTCCGTTCAAAATTA 1451
Db 379 AARACATCTACATCTTCAAGAGAGTGTAGTTATGCGAAAGTTTCCGTTCAAAATTA 438
Qy 1452 TTACCTCAATATTCACTCCAGAACGC 1477
Db 439 TTACCTCAATATTCACTCCAGAACGC 464

RESULT 10
US-10-029-386-9711/c
; Sequence 9711, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251

; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 9711
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 2.00e-67
; OTHER INFORMATION: NT HIT: g11421959, EVALUE 0.00e+00
; OTHER INFORMATION: EST_HUMAN HIT: BF237807.1, EVALUE 0.00e+00
US-10-029-386-9711

Query Match 12.6%; Score 401.4; DB 16; Length 530;
Best Local Similarity 99.8%; Pred. No. 3.7e-108;
Matches 402; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 968 ACAACGACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGCAG 1027
Db 530 ACAACGACGATTCGAGTTCGAGAGGAGCTTGGAGAAACAAATAAGGCGAGTTGTGCAG 471
Qy 1028 GCCTCTCGACAGAGAGAGAGTGCACAACTCCACGCGGAGGAGCGCCCTCCGTGGAGC 1087
Db 470 GCCTCTCGACAGAGAGAGAGTGCACAACTCCACGCGGAGGAGCGCCCTCCGTGGAGC 411
Qy 1088 CGGATCCCAAGTATCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTCCGCGCAAG 1147
Db 410 CGGATCCCAAGTATCCAGTAGCAAGAGAGAGCCACTCACTGCTCCGAGTCCGCGCAAG 351
Qy 1148 TTTTCAGAACCTACCAACGAGTGGTCTTGGACTCCAGGGTCCACAAGAGGACCGGAGGG 1207
Db 350 TTTTCAGAACCTACCAACGAGTGGTCTTGGACTCCAGGGTCCACAAGAGGACCGGAGGG 291
Qy 1208 CGGCGGCGAGTCCGCCACCATGCTGTGGAGCGGAGGAGCGCGGAGCGTGTCTCTCTG 1267
Db 290 CGGCGGCGAGTCCGCCACCATGCTGTGGAGCGGAGGAGCGCGGAGCGTGTCTCTCTG 231
Qy 1268 ACCTCGCGCCCTCTTGATGAAATGGAGCGGAGGAGGAGGAGGAGTGTGTCTGAAG 1327
Db 230 ACCTCGCGCCCTCTTGATGAAATGGAGCGGAGGAGGAGGAGTGTGTCTGAAG 171
Qy 1328 ACGATCTGAGATGGGCTTCCGAGGAGTCCCAAGGAATCCATCTCGATTA 1370
Db 170 ACGATCTGAGATGGGCTTCCGAGGAGTCCCAAGGAATCCATCTCGATTA 128

RESULT 11
US-10-029-386-23411/c
; Sequence 23411, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AECOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23411
; LENGTH: 251

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR20.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.94
; OTHER INFORMATION: EST_HUMAN HIT: W05407.1, EVALUE 1.00e-117
; OTHER INFORMATION: NT HIT: g11421959, EVALUE 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: O75362, EVALUE 1.00e-38
US-10-029-386-23411
```

```
Query Match 7.9%; Score 251; DB 16; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.5e-63;
Matches 251; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1093 CCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTCTCCGAGTGGCGCAAGCTTTC 1152
Db 251 CCCAAGTTACCCAGTAGCAAGAGAGCCCACTCACTCTCCGAGTGGCGCAAGCTTTC 192

Qy 1153 AGAACTTACCCAGTGGTCTTGACCTCCAGGGTCCACAGAAAGACCCGAGGCGCGC 1212
Db 191 AGAACCTTACCCAGTGGTCTTGACCTCCAGGGTCCACAGAAAGACCCGAGGCGCGC 132

Qy 1213 GCGGAGTCCGCCACCATCTCTGTGACGGGAGGCGCGGAGCGTGTCTCTGACCTC 1272
Db 131 GCGGAGTCCGCCACCATCTCTGTGACGGGAGGCGCGGAGCGTGTCTCTGACCTC 72

Qy 1273 GCGGCGCTCTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGGTCTTGAAGACGGA 1332
Db 71 GCGGCGCTCTGGATGAAATGGAGCCGTGGATCGAGGGAAGGTGGTCTTGAAGACGGA 12

Qy 1333 TCTGAGGATGG 1343
Db 11 TCTGAGGATGG 1
```

```
RESULT 12
US-09-783-590-9057
; Sequence 9057, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.2C1
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 9057
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (12)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (65)
```

```
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (75)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (103)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (108)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (113)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (193)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (204)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (212)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-9057

Query Match 5.1%; Score 164; DB 9; Length 267;
Best Local Similarity 90.4%; Pred. No. 1.2e-37;
Matches 226; Conservative 0; Mismatches 18; Indels 6; Gaps 5;

Qy 2315 GTAAACCCAGCCCAAGTCTGCTTTCCGGCGGAGTCCAAATCCCTGCATCTGGAGG 2374
Db 18 GTAAACCCAGCCCAAGTCTGCTTTCCGGCGGAGTCCAAATCCCTGCATCTGGAGG 76

Qy 2375 GG-AAGCAGAGCCCTCTCTGGGCGGAGGCAAGCCCTCTGA--CTTCAGGGATAGACTCTA 2431
Db 77 GGAAGCAGAGCCCTCTCTGGGCGGAGGAGGNCCTTGAACCTTCAGGGAATAGACTCTA 136

Qy 2432 GCACCTTTAGCCCCAAGT--AACCTGAAGTCCACAGACACAGAGAAATGTGGGGTCCAA 2490
Db 137 GCACCTTTAGCCCCAAGTAAACCTGAAGTCCACAGACACAGAGAAATGTGGGGTCCAA 196

Qy 2491 GGGGCGGCGGAGGCAAGCAACAGCAATCTG-AGATGTTTCTTAAACCAGTGTTCCTCCCTGC 2549
Db 197 GGGGCGGCGGAGGCAAGCAACAGCAATTTGAAGATGTTTCTTAAACCAGTGTTCCTCCCTGC 256

Qy 2550 ACCGGATAAG 2559
Db 257 AACGGATTAG 266
```

```
RESULT 13
US-10-172-118-1511
; Sequence 1511, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
```

PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 014643
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1511

Query Match 3.2%; Score 102.4; DB 17; Length 6033;
Best Local Similarity 52.4%; Pred. No. 1.9e-18;
Matches 289; Conservative 5; Mismatches 246; Indels 12; Gaps 3;
QY 379 GAAATTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCAC 438
Db 1057 GAGTTCCCGTGGAGGTGTGGCCAGGCTTTTCAGCCAGACCTGTTCTCTGAAGCGCAC 1116
QY 439 ATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAATGCGGAGAGMTTSRRS 498
Db 1117 ATGAAGAGCACCGGGCTCTTCGACACGGCTGCCACATCTCGGGCCGTAGGTTCAAG 1176
QY 499 SAGCCTTGTTCTTAAAAATCACATGCGGACACATATGGAATCGGGGCCAGAGGC 558
Db 1177 GAGCCTTGTTCTCAAGAACACATGAAGGCGACGGCCCNACAGCGGCGAGCAAGAAC 1236
QY 559 AAATGCGAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCAC 618
Db 1237 AGGCCCAAGAGTGAGCTGAGCCCCA---TCGCCACCATCAACAGTGTCCAGGAGG 1293
QY 619 GCGGCGGAGAGCATCTCTCTCTTACAAAATCTGATGTTGTTGGCTTCTTATTTCCA 678
Db 1294 GTGATCGTGGCGGCTGAGCTGAGCCCCA---TCGCCACCATCAACAGTGTCCAGGAGG 1353
QY 679 AATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAATCTTTCGGTACC 738
Db 1354 AACCTGGACAGCTTGAAGCCCAATATGTCATCCACCGAGTGCAGGC---CAGCGC 1410
QY 739 AGCAGCGCGAGACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCTCTG 798
Db 1411 ACGCGCGCCCGCGGAGGAGGGGCGGCGGCTCGGACCAACAGTCTTCTTCTC 1470
QY 799 CAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACG-----GGGAAGAGCTGTCTG 852
Db 1471 CAGTGCCTGAACTGAGGCGGTGCGCGCGCGGAGTCTCGTCCCTGGCACGAGCGCGGA 1530
QY 853 AGATGCATCCCTCAGCTGATCCGTTTCCAGGCTTCCAGGCTTGGCAGTGCCTACCAA 912
Db 1531 CGCGGGTGGCTGAGCTGACCCGCTCAACAGCTACAGGCTGCGAGTGCCTGCGCAGCGG 1590
QY 913 GGAAAGTTGCC 924
Db 1591 GGTAAGGTGGCC 1602

RESULT 14

US-10-342-887-1511
; Sequence 1511, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14

; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1511
; LENGTH: 6033
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1511

Query Match 3.2%; Score 102.4; DB 18; Length 6033;
Best Local Similarity 52.4%; Pred. No. 1.9e-18;
Matches 289; Conservative 5; Mismatches 246; Indels 12; Gaps 3;
QY 379 GAAATTAGCTGTGAGGTATGTGGGAGACATTTAGAGTCGCTTTTGTGATGTTGAGATCCAC 438
Db 1057 GAGTTCCCGTGGAGGTGTGGCCAGGCTTTTCAGCCAGACCTGTTCTCTGAAGCGCAC 1116
QY 439 ATGAGAACACACAAAGATTCTTTCACTTACGGGTGTAAATGCGGAGAGMTTSRRS 498
Db 1117 ATGAAGAGCACCGGGCTCTTCGACACGGCTGCCACATCTCGGGCCGTAGGTTCAAG 1176
QY 499 SAGCCTTGTTCTTAAAAATCACATGCGGACACATATGGAATCGGGGCCAGAGGC 558
Db 1177 GAGCCTTGTTCTCAAGAACACATGAAGGCGCACGGCCCNACAGCGGCGAGCAAGAAC 1236
QY 559 AAATGCGAGCAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGTCTGTCAGGTGCAC 618
Db 1237 AGGCCCAAGAGTGAGCTGAGCCCCA---TCGCCACCATCAACAGTGTCCAGGAGG 1293
QY 619 GCGGCGGAGAGCATCTCTCTCTTACAAAATCTGATGTTGTTGGCTTCTTATTTCCA 678
Db 1294 GTGATCGTGGCGGCTGAGCTGAGCCCCA---TCGCCACCATCAACAGTGTCCAGGAGG 1353
QY 679 AATAAGAAAGTCTAATTGAGCACCGCAAGGTGCACACCAAAAATCTTTCGGTACC 738
Db 1354 AACCTGGACAGCTTGAAGCCCAATATGTCATCCACCGAGTGCAGGC---CAGCGC 1410
QY 739 AGCAGCGCGAGACAGACTCTCCACAGGAGGAATGCCGTCTCGAGGGAGGACTTCTCTG 798
Db 1411 ACGCGCGCCCGCGGAGGAGGGGCGGCGGCTCGGACCAACAGTCTTCTTCTC 1470
QY 799 CAGTTGTTCAACTTGAGACCAAAATCTCACCTGAAACG-----GGGAAGAGCTGTCTG 852
Db 1471 CAGTGCCTGAACTGAGGCGGTGCGCGCGCGGAGTCTCGTCCCTGGCACGAGCGCGGA 1530
QY 853 AGATGCATCCCTCAGCTGATCCGTTTCCAGGCTTCCAGGCTTGGCAGTGCCTACCAA 912
Db 1531 CGCGGGTGGCTGAGCTGACCCGCTCAACAGCTACAGGCTGCGAGTGCCTGCGCAGCGG 1590
QY 913 GGAAAGTTGCC 924
Db 1591 GGTAAGGTGGCC 1602

RESULT 15

US-10-074-475-93
; Sequence 93, Application US/10074475
; Publication No. US20030092898A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Hu, Ping
; APPLICANT: Recipon, Herve
; APPLICANT: Karra, Kalpana
; APPLICANT: Cafferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chonghua
; TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific
; FILE REFERENCE: DEX-0313
; CURRENT APPLICATION NUMBER: US/10/074,475
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/268,292
; PRIOR FILING DATE: 2001-02-13
; NUMBER OF SEQ ID NOS: 295


```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 93
; LENGTH: 8156
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-074-475-93
```

```
Query Match      3.2%; Score 102.4; DB 14; Length 8156;
Best Local Similarity 52.4%; Pred. No. 2.3e-18;
Matches 289; Conservative 5; Mismatches 246; Indels 12; Gaps 3;

Qy 379 GAATTTAGCTGTGAGGTATGTGGCAGACATTTAGAGTCGCTTTTGATGTTGAGATCCAC 438
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1057 GAGTTCCCGTGGAGGTGTGTGGCAGGCCCTTCAGCCAGACCTGGTTCCTGAAGGCGCAC 1116

Qy 439 ATGAGACACACAAGATTCTTCTACCTTACGGGTGTACATGTGCGGAGAGGTTTSPRS 498
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1117 ATGAAGAAGCACCGGGGCTCTTCGACCACCGTGCACATCTGCGGCCGTAGGTTCAAG 1176

Qy 499 SAGCCTTGGTTTCTTAAAAATCACATGCGGACACATATGGCAATCGGGGCCAGAGC 558
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1177 GAGCCTTGTTCTTCAAGAACCAACATGAAGGGCAGCGGCCCAAGACGGGCGAGCAAGAC 1236

Qy 559 AAATGCGACGAAGGCTTGGAGAGTAGTCCAGCAACGATCAACGAGGTGCTCCAGGTGCAC 618
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1237 AGGCCCAAGAGTAGCTGGACCCCA---TCGCCACCATCAACAACGTTGTCAGGAGGAG 1293

Qy 619 CGGGCGGAGAGCATCTCTCTCTTACAAATCTGCATGTTTGTGGCTTCTATTTCGA 678
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1294 GTGATCGTCGGCGGCTGAGGCTCTACGAGGTCTGCGCCAAGTCGGGAACCTGTTTACA 1353

Qy 679 AATAAGAAAGTCTAATTGAGCAGCGCAAGGTGCACACCAAAAAAACTGCTTTGCGTACC 738
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1354 AACCTGGACAGCTTGAACGCCCAACATGCCATCCACCGCAGATCGAGGC---CAGCCGC 1410

Qy 739 AGCAGCGCGCAGACAGACTCTCCACAGGAGGAATGCCGTCTCTGAGGGAGGACTTCTG 798
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1411 ACGCGCGCCCGCGGAGGAGGGGCGGAGGGGCCCTCGGACACCAAGCAGTCTTTCTC 1470

Qy 799 CAGTTGTTCAACTTGAGACCAAAATCTCACCTGNAACG-----GGGAAGAAGCCTGTC 852
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1471 CAGTGCCTGAACCTGAGGCCGTTCGGCGCGCGCGACTCGTGCCTTGGCACGAGGCCGGA 1530

Qy 853 AGATGCATCCCTCAGCTCGATCCGTTCCACACCTTCCAGGCTTGGCAGCTGGCTACCAAA 912
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1531 CGGCGGGTGGCTGAGCTGGAACCGGTCAACAGCTACCAAGGCTTGGCAGCTGGCCACGCGG 1590

Qy 913 GGAAAAGTTGCC 924
Db ||| ||| ||| |||
1591 GGTAAGTGGCC 1602
```

Search completed: July 4, 2005, 11:51:05
Job time : 1863.96 secs